# Y-STR HAPLOTYPE DIVERSITY AND mtDNA HVI SEQUENCE VARIATION AMONG MAJOR TRIBES OF CHARSADA AND MARDAN DISTRICT

**SADIA TABASSUM**

**DEPARTMENT OF GENETICS HAZARA UNIVERSITY MANSEHRA 2016**

**HAZARA UNIVERSITY MANSEHRA**


## Department of Genetics

# Y-STR HAPLOTYPE DIVERSITY AND mtDNA HVI SEQUENCE VARIATION AMONG MAJOR TRIBES OF CHARSADA AND MARDAN DISTRICT

**BY**

**Sadia Tabassum**

This research study has been conducted and reported as partial fulfillment of the requirements of PhD degree in Genetics awarded by Hazara University Mansehra, Pakistan

Mansehra,

The Monday 11, April 2016

**Y-STR HAPLOTYPE DIVERSITY AND mtDNA HVI SEQUENCE VARIATION AMONG MAJOR TRIBES OF CHARSADA AND MARDAN DISTRICT**


Submitted by

**Sadia Tabassum**
PhD Scholar

Research Supervisor

**PROF. DR. HABIB AHMAD**
Chairman
Department of Genetics,
Hazara University Mansehra

Co Supervisor

**DR. MUHAMMAD SHAHID NADEEM**

Department of Genetics Hazara
University Mansehra


# DEPARTMENT OF GENETICS
# HAZARA UNIVERSITY MANSEHRA
# 2016


بسم الله الرحمن الرحيم

DEDICATED

T

MY RESPECTED PARENTS

Who made me what I am today

(Ma Th Live Long)

AN

**MY DAUGHTER FATIMA**

**ACKNOWLEDGMENTS**

All the praise and glory is for almighty Allah who has showered on me his countless blessings and unbeatable strength to cope up with the challenges of life, without which the journey of life was impossible to go through.

This work was carried out during the years 2012 - 2016 in the joint venture of Human Genetics lab, Hazara University Mansehra and Molecular Anthropology lab, University of Illinois at Urbana Champaign, USA. After the intense period of 4 years, today is the day writing this note of thanks is the finishing touch of my thesis. It has been a period of intense learning for me, not only in the scientific arena, but also on a personal level. Writing this thesis has had a big impact on me. I would like to reflect on the people who have supported and helped me so much throughout this period.

I owe my deepest gratitude to my supervisor Professor Dr. Habib Ahmad (*TI*) the respected vice Chancellor of Hazara University Mansehra. Without his optimism in this project, enthusiasm, encouragement and support, this study would hardly have been completed. I could not have imagined having a better advisor and mentor for my PhD study. He consoled me in every difficult and challenging situation and boosted up my morale and provided me a conducive and congenial working environment in the lab. He actually had triggered and nourished my intellectual maturity and equipped me with knowledge that I will benefit from, for a long time to come. It was an extraordinary situation for me to be a part of his research team. It would not be an exaggeration to say that

I came to know what actually science means in a true sense during my stay with Professor Dr. Habib Ahmad (*TI*).

Beside my supervisor, I would like to thank rest of my thesis committee members for their insightful comments and encouragements.

I also express my warmest gratitude to my co supervisor Dr. Shahid Nadeem, Assistant Professor Department of Genetics, Hazara University Mansehra who has made available his moral support in number of ways.

This study was sponsored by Higher Education Commission of Pakistan under "5000 Indigenous fellowship program Batch V", as well as International Research Support and Initiative Program (IRSIP) and I was given opportunity to go to the Department of Molecular Anthropology, University of Illinois at Urbana Champaign USA to carry out my research work to achieve my proposed targets in the world top ranked university, where I could improve my professional skills. It was wonderful experience of my life.

I am deeply grateful to Prof. Dr. Ripan S. Malhi, Department of Anthropology, UIUC, USA, for making it possible to carry out my experimentation in his lab and for providing me an opportunity to be a part of his research team. I worked under his kind supervision and learned a lot. It was the wonderful and unforgettable experience of my life to be among the world class scientists having advanced laboratory, equipped with modern and sophisticated

scientific instruments. I found Professor Ripan S. Malhi as the most cooperative and nice man. He guided me throughout during my stay there in USA. Besides helping me out in my research, he gave me the permission to attend his lectures on Molecular Genetics and Anthropology in the University for the whole period I spent there. It was an extraordinary situation for me to be among the American students and attending lectures of the world renowned scientist.

I would like to show my gratitude to Ms. Amanda Owings, Kelsey Witt and Dr. Hongjie Lee, for their support, guidance and company. I learned many basic techniques used in the laboratory from them.

I am indebted to my all colleagues and fellows in Human genetics lab and department of Zoology to support me, who always boosted my self-confidence during achievement of my goal. I am grateful to the sampling team, especially Mr. Muhammad Ismail Khan, Mr. Rehman ullah, Mr. Muhammad Tariq, Mr. Inam ullah and other team members, without their support this research was impossible. I pay especial thanks to my dear friend Samina Yasmin for always being good companion and supportive friend.

I would like to show my gratitude to Mr. Asif for his encouragement and excellent timely cooperation and providing me a logistic support without which it was difficult to accomplish this task.

I owe my profound gratitude to Dr. Muhammad Ilyas, Assistant Professor, Department of Genetics for his fruitful suggestions in write-up.

Last but not least I would to thank my family for supporting me spiritually throughout my study and my life in general. No words of acknowledgements can be ample to express magnificent love and lavish cooperation of my loving parents who put the cornerstone to my learning character, showing me the joy of intellectual pursuit ever since I was a child. I would also like to thank them all for their wise counsel and sympathetic ear. Beside all, I would like to thank my little angel Fatima for being my strength and who has been a driving force behind the entire struggle made.

**SADIA TABASSUM**

# TABLE OF CONTENTS

## LIST OF TABLES

# LIST OF FIGURES

# ABSTRACT

Substantial genetic diversity exists among the people of Pakistan and understanding evolution of this diversity is complicated due to several waves of migration from populations in the North and Northwest. Pathans, one of the largest ethnic groups of Pakistan residing mostly in the northwestern part which remained unexplored with respect to their biological affinity and haplotypic diversity, is the subject of this exploration.This dissertation reports the results of the PhD project wherein, we have assessed the extent of genetic diversity using Y-STR profiling and *mt*DNA sequence analyses of HVI control region of Pathan populations of Mardan and Charsada districts of KP province, Pakistan.Sum of 374 buccal swabs were collected from five major population of the two adjoining districts viz Charsada and Mardan. Y-STR profiling was carriedout using Promega PowerPlex® Y23 System. Mitochondrial HVI data was also generated for 165samples. Y-chromosomal and *mt*DNA haplogroups were assigned to each sample using a haplogroup predictor and phylotree, respectively. Principle Coordinate Analysis (PCoA) plots were generated by combining our data set with other published datasets from neighboring populations of central Asia, Middle East, Europe and South Asia. Our results revealed that the most frequent Y-STR haplogroup found was R1a (49.2%), followed by G2a (17.9%) and L (9.6%), while *mt*DNA HVI macro haplogroups R (63.4%), M (26.8%) and N (8.6%) were observed more frequent among Pathans of the area. Some novel mitochondrial haplogroups (*mt*Hgs) of M and R subclades have also been recorded for the first time for Pathans. Analysis of our results

revealed that no significant genetic substructure was available among all the five populations with respect to 23 Y-STRs. Furthermore the Pathans of Charsada and Mardan are clustered together with little genetic difference. These people show homology with already studied Pathans from Pakistan, the Yousafzai tribe from northern proximity of Pakistan, as well as with the Pathans from Afghanistan and with some populations from Russian Federation, when compared on the basis of Y-STR profiles. The clustering pattern was different when compared on the basis of mitochondrial HVI data sets of different neighboring populations, which depicts more male mediated gene flow among populations than female mediated gene flow. In addition, the mitochondrial gene pool of these populations seems to be influenced by the Turkish invasion. The gene pool of the Pathan populations of Mardan and Charsada districts exhibit genetic affinities with other Pathans from Pakistan and Afghanistan, and exhibit shared haplogroups with Middle Eastern, Caucasian, South Asian and Central Asian people. The study provides a sound reference for collating molecular anthropology and genetic structure of the people of Pakistan.

# Chapter 1
# INTRODUCTION

A comprehensive rational interest, about the origin of mankind has always been compelled the human brain to seek logical answers. In fact the key inspiration behind this grilling and the search for human genetic variation has always been a challenge to cope up with and to explore the population structure in general and genetics of diseases in particular. Genetic variations available in the modern people also reflect the past of mankind, its origin and the spread of anatomically modern humans and their demographic history. It is increasingly more evident that knowledge about these processes is equally indispensable for a proper understanding of the genetics of complex diseases.

Decades long research of genetic variation has revealed that a significant degree of diversity and variation lies both within and between existing human populations around the globe. Despite the enormous amount of data gathered during this relatively long time span of "classical era", has highlighted many basic problems but left them largely unresolved. Increasingly more informative "DNA era", rapidly expanding during the last 35 years, took up the same list of problems and is formulating new ones. Largely irrespective of what kind of general questions related to demographic history are being asked, the present-day genetics investigates variation in three different systems i.e. bi-parental autosomal chromosomes, paternally inherited Y chromosome and maternally inherited mitochondrial DNA. In spite of having manifold large size of autosomal genes

than that of Y chromosome and mtDNA, the uniparentally inherited markers have powerful advantages, making them the favorite and handy tools of choice for the population geneticists. Besides their uniparental mode of inheritance, the advantage of their lacking of recombination is of prime importance that helps them to reconstruct genetic lineages back to the most recent common ancestors (MRCA). i.e. The Y chromosomal Adam and mitochondrial Eve. Furthermore in combination with information about the variation in autosomal genes, a promising synthesis is hopefully possible in the future. My thesis is about Y chromosomal STR haplotype and mtDNA HVI variation in major Pathan populations (tribes) of the two historically important geographically adjacent areas that remained under the influence of many foreign invaders. In order to understand the Y chromosomal STR variation and mtDNA HVI variation and its origin in these populations, additional background knowledge of those living outside this area is needed. Due to their scrupulous properties Y-STRs mtDNA allow to construct individual genealogies and connect them to a global phylogenetic tree of all humans. Basal branches of these Y Haplogroup and mtDNA trees have been shown to be highly continent specific, making these uniparentally inherited loci more valuable makers compared to other tools where their overall variation is hidden mainly within populations. The study of the spread and variability of Y-haplogroups and mtDNA clusters enable the scientists to address questions related to the early population movements of anatomically modern humans.

Through the collective effort of many labs and researchers around the globe, the Ychromosomal and mitochondrial DNA clusters (haplogroups) have been investigated and explored in different continental regions.

South Asia comprising India, Pakistan, countries in the sub-Himalayan region and Myanmar was one of the first geographical regions to have been inhabited by modern humans. This region has served as a major route of dispersal to other geographical regions, including Southeast Asia. Populations of Pakistan show high genetic differentiation and extensive structuring as a result of evolutionary antiquity and commonly practiced endogamy. Linguistic variations of populations present the best account of genetic differences observed in this region of the world.

The progression of local population expansion and consequent dispersal to newer areas, coupled with evolution of cultural practices and geographical distance acting as a barrier to human contact, also resulted in the formation of intra-marrying groups; groups within which there was considerable gene exchange among individuals, but between which there was little genetic exchange. The development and evolution of language and culture embellished the formation of such groups.

Genetic diversity within a geographical area is a logical indicator of the age of the population" s habitation in that specific area. However, it is also dependent on the population" s effective sizes, thus entailing that the assessment of antiquity may not be simple. It is now well-established that Africa exhibits the highest genetic diversity among continental populations. The genetic diversity in Indian subcontinent, "the heartland of

south Asia", is second to that of Africa. Mitochondrial DNA (mtDNA) and the non-recombining region of the Y-chromosome (NRY) based Molecular genetic studies have supported a single southern dispersal route to Indian sub continent. The populations of Indian sub continent including Pakistan and India are all derivatives of mitochondrial M and N lineages, which themselves are the derivation of L3 lineage that now found only in Africa. The M and N lineages probably diverged from L3 shortly after their dispersal from Africa (Quintana *et al*., 1999). Similarly Y-chromosomal data is also indicative of southern dispersal rout and this is supported by the presence of C and D Y-haplogroups only in the Asian continent and Oceania (Kivisild *et al*., 2003; Underhill *et al*., 2001) and not in western Eurasia and North Africa.

Afghanistan being a landlocked nation in southwest Central Asia has central highlands of Hindu Kush Mountains extending from the northeast of the country to the southwest, separating Afghanistan into northern and southern provinces. Being at the intersection of Central Asia, South Asia, and the Middle East, Afghanistan has served as crossroads for human migrations and pilgrimages, including an important stop along the Silk Road. Mesolithic artifacts, Neolithic pottery of about 7.2kya, bones of domesticated animal and tools i.e. the sickle blades used to collect wild grasses, have been discovered during excavations in the Ghar-i-Mar site in north Afghanistan. These discoveries supported the early cultivation of wheat and barley around 9–11kya and domestication of animals around 7–9kya (Dupree *et al.*, 1972). Other recent archeological findings include Buddhist artifacts transported northward from India along the Silk Road, as well as captions and inscriptions engraved on rocks in ancient Hebrew dating from the eleventh to the

thirteenth centuries (Katzir, 2001). Historians speculated the onset of urbanization in Afghanistan around 4 to 4.5kya (Runion, 2007). According to Historical records Aryans from Iran were the first to occupy Afghanistan in eighth century B.C.E.), followed by Persians in sixth century B.C.E and Greeks in fourth century B.C.E. Later Mauryans from India and GrecoBactrian escorted Buddhism into this region while s Arabs and Mongols launched Islam (Runion, 2007). Dari and Pashto are two languages of Indo-European origin spoken by two major Ethnic groups of Afghanistan; the Tajiks and Pathans or Pathans, respectively. Pathans comprising the most prevalent ethnic group in Afghanistan (42%), inhabit primarily the south of the Hindu Kush mountains. According to legendary record and oral traditions, these

Pathans are considered to be the ancestors of Pathans in the Pakistan and Northern India. Different origins for Pathans, including Greek and Jewish ancestry, have been suggested (Runion, 2007; Caroe, 1976). These connections have not been confirmed and a few efforts characterizing the genetic structure of this group in Pakistan (Mohyuddin *et al.*, 2001; Qamar *et al.*, 2002; Mansoor *et al.*, 2004; Sengupta *et al.*, 2006; Firasat *et al.*, 2007), Afghanistan (Di Cristofaro, 2012) and India (Noor *et al.*, 2009) have been made. As previous studies were carried out without the tribal and subtribal status of the Pashtuns, so such study was of immense need to segregate them on the basis of their subtribal and caste level.

The present study was undertaken to ascertain, for the first time, the genetic diversity of major Pathan tribes of Pakistan residing in Charsada and Mardan district of KP province

such as Yousafzai, Mohmand, Muhammadzai and Kakakhel Mian, utilizing Ychromosomal STR profile and mtDNA HVI sequence variation. The data produced were subsequently compared with formerly published geographically and ethnically connected indigenous and global populations to explore paternal and maternal signals of modern human" s dispersals across this study area. In addition, we assessed genetic affinities between Pathans from our study area and Pathans from neighboring Afghanistan as being descendants of the same Y-chromosome, as well as their hypothesized phylogenetic relation to Greek and Jewish populations.

Based on the legendry information, it was hypothesized that the populations under investigation are the descendants and immigrants of Pathan ethnic group of Southern Afghanistan. So they must have close genetic ties with their suspected ancestors. On the basis of above mentioned hypothesis, several following queries were raised to be justified after this research.

1. What clusters (Haplogroups) are spread among populations of Charsada and Mardan?
2. How do these clusters connect to the global tree?
3. What is the diversity (%) of Y haplogroups and mtDNA haplogroups in these populations as compared to the populations who had invaded these areas and ruled for a long time?
4. What has been an impact of succeeding gene flow between these populations?

5. What has been a contribution of invading populations to the maternal and paternal gene pool of the populations under study?

The central focus of the literature overview of the current thesis will be therefore an overview of out of Africa migration and colonization of South Asian land mass by modern humans, the legendary ancestry of the populations under investigation and history of study area and finally the general skeleton tree connecting and characterizing the Y haplogroups and mtDNA haplogroups worldwide generally and in south Asia specifically.

# Chapter2
# REVIEW OF LITERATURE

Charles Darwin was the pioneer in suggesting the common African ancestry of human beings (Lafrenier, 2010). On the basis of some shared attributes he further proposed African apes to be their ancestors (Bowler *et al*., 2003). Later on his proposition was rejected and after several genome analyses of people from different racial backgrounds, it was concluded that all modern humans evolved from one 'lucky mother'
in Africa about 150,000 years ago (Cann *et al.*, 1987; Vigilant *et al.*,1991). The recurrent development in molecular genetics and availability of analytical tools has revolutionized molecular anthropology which needs to be strengthened in Pakistan
also.

## 2.1 Colonization of South Asia

According to Darwin human population started dispersing around 125,000 years before present (BP) using one route from the Nile Valley directed towards the Middle East, around 120,000–100,000 years BP and the second one through Bab-el-Mandeb Strait on the Red Sea, crossing through the Arabian Peninsula and finally heading towards the present-day United Arab Emirates around125,000 years BP and Oman around 106,000 years BP (Lafrenier, 2010) and then consequently reaching the Indian Subcontinent atJwalapuramaround75,000 years BP. Despite lacking any fossil record, the similarity among the stone tools collected from all these places, suggested that the modern humans to be their creator (Bower, 2011). These findings might give some support to the claim

that modern humans from Africa arrived at southern China about 100,000 years BP *(*Liu *et al*., 2010), and the Liujiang hominid controversially dated at 139,000– 111,000 years BP (Shena *et al., 2002*). Dating results of the Lunadong (Buding Basin, Guangxi, southern China) teeth, which include a right upper second molar and a left lower second molar, indicate that the molars may be as old as 126,000 years.

For the previous exoduses from Africa genetic analyses based on the Y chromosome and mitochondrial genome doesn" t provide any clue, as they represent only a small part of the human genetic material.It could be assumed that those modern humans became extinct because of Toba catastrophe around 74,000 years BP. However, according to some arguments there was no remarkable impact of this catastrophe on human population (Balter, 2010).

Someother authorities claim that *Homo sapiens* first appeared around 200,000 years ago in Ethiopia (White *et al*., 2003). After reaching Near East around 125,000 years ago, they moved back to Africa, as their settlements were replaced by Neanderthals. Now days it is strongly believed that the first modern human dispersal across Asia started about 75,000 years ago across the Bab el Mandib linking Ethiopia and Yemen. From there, some of these people moved further to South Asia around 50,000 years ago, and then to Australia by 46,000 years ago(Bowler*et al*., 2003). *H. sapiens* reached Europe around 43,000 years ago (Wilford, 2011), replacing the Neanderthals around 24,000 years ago. They reached East Asia some 30,000 years ago.

Migration of several languages and cultures has also been noticed with these human migrations. The initial exodus out of Africa 125,000 years ago was followed by several waves of migration via the Arabian Peninsula into Eurasia around 60,000 years ago, where one group of migrants rapidly passes through coastal areas of the Indian Ocean and other group migrating north towards Central Asia.

Genetic evidence proved that modern humans have gone through genetic bottleneck resulting in reduced genetic diversity and then gone through a dramatic population growth from geographically dispersed populations around 50,000 years ago. And this idea of bottleneck was also supported by Henry Harpending" s findings as well as geological and climatological evidences from Lake Toba explosion (Harpending *et al.*, 2009), that had resulted in overall population reduction up to 15,000 individuals that consequently leading to rapid racial differentiation as a possible consequence of founder effect and increased genetic drift (Ambrose, 1998).

According to Some genetic evidences the migration out of Africa had occurred along two different routes. However, other investigations revealed that a single migration occurred, followed by swift northern migration of a subset of the group. The group who followed southern route from West Asia spread generation by generation around the coast of Arabia and Persiafinally reached Indo-Pak subcontinent and headed towards Australia around 55,000 to 40,000 years ago (fig.1). The group directed towards north (East Asians were the second group) ventured inland (Maca *et al*., 2001) and exuded to

Europe, ultimately displacing the Neanderthals. They also radiated to India from Central Asia (Bowler *et al*., 2003). Some recent evidences suggest that the human migration from Africa to Southeast Asia and Australia have occurred between 80,000 and 120,000 years ago (Ewen, 2015) following multiple routesand the *H. sapiens* interbred with other species like Neanderthals in Europe, Denisovans in Central Asia and Homo erectus (Dannel, 2012).

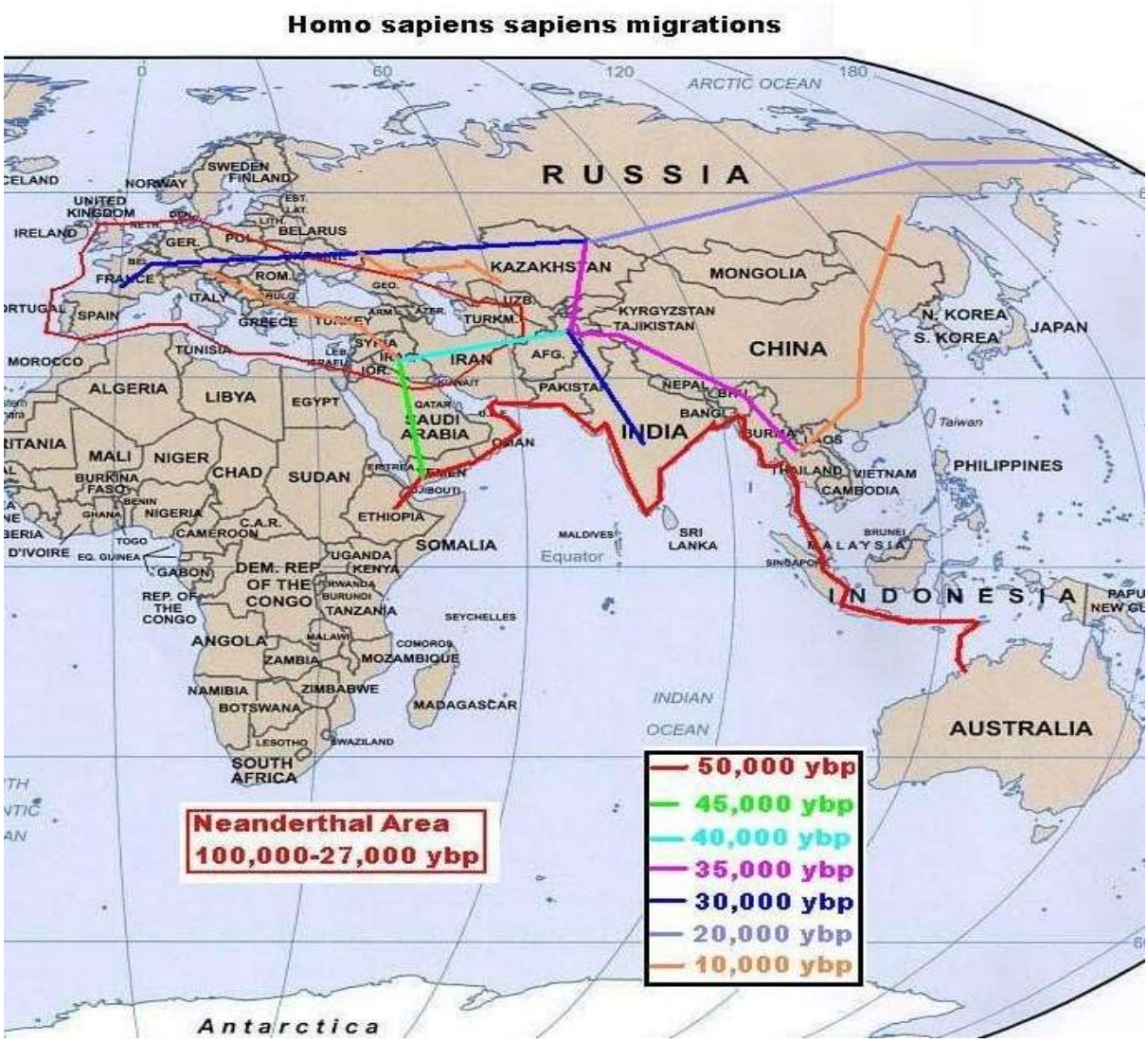



Fig.1. Probable human out of Africa migration time scale. (Source:

http://www.roperld.com/YBiallelicHaplogroups.htm#routes)

**2.2 Pakistani civilization**

Pakistan is diverse country with variety of cultures and ethnic standings available in its more than 180 million people. Historically it has been an important trading hub since the ancient Indus valley and Gandhara civilizations, Pakistan has had been exposed to many invaders and traders from around the world i.e. Africa, Middle East, Caucasus, Mediterranean and China. This influx has contributed to the vibrant mix of people in modern day Pakistan. The majority of Pakistani population belongs to an IndoEuropean ethnic groups comprisingof various smaller ethnic grouping. Punjabis being the largest group constitute almost 44% of the population, Pathans being second largest ethnic group constitute 15%, Sindhi constitute 14%, Saraiki 10%, Muhajir 7%, Baluchi
3% and the remaining 4% of the population consist of Hindkowans, Gujjars, Kashmiri, Potoharis, Chitralis, Hunzakut, Dards, Balti and several other small groups.

**2.3 The Pathans**

Pathans, the Pashto speaking people also spelled as Pashtuns, Pakhtuns, Pukhtuns, broadly referred to as Pathans or ethnic Afghans (Banuazizi *et al*., 1994) belong to an Eastern Iranian ethno-linguistic group residing for the most part of Eastern and southern Afghanistanand in Khyber Pakhtunkhwa Province, Federally Administered Tribal Areas and parts ofBaluchistan Province of Pakistan. They are characterized by Pashto language, their culture, honor and the pre-Islamic code of conduct Pukhtunwali they practice

(Oliver, 1980).Their unified modern past is sometimes referred to the establishment of the rise of the Durrani Empire in 1747.

Majority of Pathan tribes live in an area stretching from western Pakistan to southwestern Afghanistan. Additional Pathan communities reside in the Northern Areas, Azad Kashmir and Sindh provinces of Pakistan as well as throughout Afghanistan and in the Iranian region of Khorasan. A large migrant worker community lives in the countries of the Arabian Peninsula and in smaller communities in Europe and North America. A sizable community of largely putative ancestry
calls India home. Important metropolitan centers of Pathan culture include Mardan and Kandahar. In addition, Peshawar, Quetta and Kabul are ethnically mixed cities with large Pathan populations. With 1.5 million ethnic Pathans, Karachi stands the largest Pathan city in the world.

Pathans with total population of over 15.4 percent or 25.6 million people are widely dispersed in Pakistan inhabiting vast area from Khyber to Bolan including Peshawar, Quetta and Karachi (the largest Pathan city in the world) (Shehzad, 2007).They are also found in vast majority in Afghanistan with the estimated population of 42% or 12.5 million people, in Khorasan (Iran) and India (Lewis, 2009).

### 2.3.1 History and origin

The Pathan areas have faced many foreign invasions of Aryans, Medes, Persians, Mauryans, Scythians, Kushans, Greeks, Haphthalites, Turks, Mongols and Arabs. They

have an ancient history, much still unresearched. Various scholars and researchers have proposed different theories about the origin of Pathans.

### 2.3.2 Historical references

Herodotus, the famous Greek historian first referred to the people called Pactyans, inhabiting Arachosia, the eastern frontier of the Persian Satrapy in the 1st millennium B.C.E. (Rawlinson, 1858). Another reference about the existence of a tribe called Pakhtas residing in the region of Pakhat in Afghanistan, is given in the Rig-Veda, and some historians have hypothesized them to be the early ancestors of the Pathans. Some others linked Bactrians to the Pathans who used other similar Middle Iranian dialect.

Historically Pathans are referred to as ethnic Afghans. Theterm Pathan was used as synonym of Afghan until the dawn of modern Afghanistan and the establishment of Durand Line, a 250 Km long border line drawn by the British Mortimer Durand in 1893. According to various scholars it was believed, "The term Afghan was first used in 982 C.E." in Hudud-al-Alam (Minorsky, 1940). The Pathans consider Afghans to be their common renowned ancestor. According to Some historians Pathans believed to emerge in the region of Kandahar and the Suleiman Mountains, and started expanding 1000s of years ago. At the site of their origin they would used to be in close proximity with the Persians and Mauryans and may have been following Zoroastrians,

Buddhism, Hinduism or Judaism before to the advent of Arab Muslims in the seventh century (Dupree, 1977; Holdich, T. H. 1899).

### 2.3.3 Anthropology and linguistics

The Pathans have been considered as the modern day descendants of Scythians speaking Pashto, a sub-branch of the Indo-Iranian branch of the Indo-European family of languages. Pashto language has ancient origin and has similarity with Bactrian and Avestan language. Pamir dialects, such as Ossetic, Shughni and Wakhi are considered to be its closest modern relatives. Pashto has borrowed many words from other languages like Sanskrit and Persian. Foreign Invaders also have left their language foot prints on Pashto and it has borrowed many words from Arabic, Greek and Turkic languages (Awde *et al.*, 2002).

According to Yu V. Gankovsky, the Pathans emerged as a "union of East-Iranian tribes that provided the ethnic platform for the Pathan ethno genesis, in the middle of the first millennium C.E. (Gankovsky, 1982). Early precursors to the Pathans had beensothe ancient Iranian tribes that stretched all through the eastern Iranian Plateau (Harvey, S. 2014). Based on the Southern or Northern Pashtodialect they speak, the Pathans either called Pathans or Pukhtuns respectively. Phenotypically, Pathans generally represent the Mediterranean people (Blood, 2001) having light hair tones and eye colors. Several Anthropologists documented the oral traditions about the origin of Pathans. Some anthropologists lend credence to the mythical oral traditions of the Pathan tribes

themselves. For instance, the Encyclopedia of Islam, highlighted the theory of Pathan descent from Israelites which could be traced from "Maghzan-e-Afghani" that was compiled in 17th century C.E. by a famous scholar in the court of Mughal emperor Jahangir, named Naimat ullah Haravi. Several other Pathan writers and historians followed the same theory in their writings. Same theory was also referred by Olaf Caroe in his writing "The Pathans".

According to the scriptures of sir William Johns, Afghans are considered to be one of the lost tribes of Bani Israel who broke away from captivity of Babylonian king Bakht Nasr and took refuge in the area of Asarah in Afghanistan. Armia and Barkhia were the two of the six sons of Sawal or Saul, a successor of Yahuda, a son of Hazrat Yaqub (AWS). Armia had a son Afghan and Barkhia had a son named Asif. Afghan was selected as a commander of an army in the kingdom of Dawood and Suleiman, after the death of Sawal. Later when Palestine was captured by Bakht Nasr, Afghan and Asif migrated to Ghor and took control of the area from the local tribes. Similarly, according to *Taaqati-Nasiri,*people called Bani Israel settled in Ghor, Afghanistan, and then started migrating towards south and east. It is documented that in 722 BCE after the conquest of Neo-Assyrian empire, the descendants of Afghan and Asif who were expelled from their homeland Palestine and lost their way, were ten in number. One of the lost tribe moved towards Mecca and met an Islamic worrier and an Israelite tribes man, Khalid bin Waleed. He then sent a message to the Afghans settled in Ghor to embrace Islam. A delegation under the command of Qais moved to Madina where they met Prophet

Muhammad (PBUH) and embraced a new faith of Islam and Qais was given a new name "Abdur Rashid" by The Prophet. Later he was given a title of "Bathan" where B was replaced by P with the passage of time and all his descendants were called Bathans or Pathans.

Those Bani Israel references are in good promise with the traditional view held by Pathans, that the tribe of Joseph among other Hebrew tribes inhabited this region when twelve tribes of Israel from twelve sons of Jacob dispersed. Bani-Israel theory stated in Maghzan-e-Afghani has been interrogated due to lack of historical justifications and linguistic inconsistencies, for examples the exiled of ten lost tribes from Assyria. Whilein *Maghzan-e-Afghani* it is referred to the ruler of Persia, authorizing those tribes to go east to Afghanistan (Harawi *et al*., 1960). That contradiction can be enlightened by Persian control over the Assyrian Empire" s lands after taking control over Babylonia. Yet there is no record for mobility of these Israelites to further east, provided by ancient authors yet, that oral tradition has been widely accepted among
Pathans. The Pathans residence in the region of Afghanistan has also been mentioned in Rig Veda that was probably composed before 1200 B.C.There is no evidence of Israelite or Jewish connection provided by ancient authors before the conversion of the Pathans to Islam.

Another group of historians believe in the Aryan descent of Pathans. According to some historians, Pathans are considered to be of Northern European lineage. Some others traced their ancestry to the Southern Russia while some consider them to be of Mongolian

and Chinese Turkistani origin. A vast group of modern researchers consider the area of "Bakhtar" between Pamir and Oxus River to be their possible birth place.

After a gradual increase in number, they started moving out of this area and scattered towards the valley of Swat and Indus River after crossing Hindukush. This group of people was called Indo-Aryana. They kept on moving towards Punjab and finally reached the valleys of Ganga and Jumna overthrowing the local Dravidian tribes. Another group from Bakhtar migrated westward towards Iran. Some historians consider them to be the ancestors of Pathans. The modern researches like Faster Tytler proposed a "Mixed Race Theory"which documented the Aryan origin of Pathans having elements of Mongol, Turkish and other strains. This idea was strengthened by the opinion of Charles Miler about the amalgamation of Greek, Mongol, Persians, Scythians and Turks to the afghan stock. Several vestiges and leftovers of Syriac language were found during mining and excavation of the Gandhara, Kandhar, Taxila and Lughman areas, providing strong evidence of the presence of Syriac people in the area. Later in 5$^{th}$ centaury this land was invaded by Haphthalites, "white Huns". In Umayyad supremacy, Arabs invaded the area and merged with the local population. In 13$^{th}$ century the "yellow race" of Changez Khan also steeped in and got mixed up with the local Afghan population.

Some historians connect the Pathan or Afghan lineage from Hazrat Ibrahim through his six sons who settled in the North West area of Iran from where they were expelled and inhabited "Parthia" or Pasht or Pakht. For which they were later called "Pashtuns" (Kakakhel, 1981).

These mythical oral traditions have been strengthened as a consequence of long cultural and political struggle between Pathans and the Mughals, elucidating the historical background for the formulation of the myth, the inconsistencies of the mythology, and the linguistic investigation that disproves any Semitic origins (Harawi *et al.,* 1960). Several other Pathan tribes claim to be descended from Arabs, including some i.e. Sayyid or Kakakhel Mian even claim their descentfrom the Muslim

Prophet Muhammad (PBUH). Some other Pathan groups i.e. Afridi, Khattak and Sadozai from Peshawar and Kandahar also claim their descent from Alexander the Great of Greece (Mansoor *et al*., 2004).

**2.3.4 Ancestral alias**

Pathan can only claim to be Pashtun if his father is Pashtun. The patrilineal definition is based on an important conventional law of Pashtunwali. That law has maintained the tradition of exclusively patriarchal tribal lineage intact. Under such definition, an ethnic Pathans have less concern with the language he speaks and more for his father. So, the Pathans who have lost both the language and seemingly many of the ways of their alleged ancestors can remain "Pashtun" by tracing their fathers' ethnic heritage back to the Pashtun tribes as they are said to be descended from Qais Abdur Rashid who has been considered as a possible Progenitor of Pathans (Allah *et al*., 2013). According to ancient historians, Qais, travelled to Arabian Peninsula to meet Prophet Muhammad in Madina and after embracing Islam returned back to Afghanistan and present day

Pakistan. Then got married to Khalid bin waleed" s daughter and ostensibly had many children and four sons Sarban, Baitan, Ghourghusht and Karlan, who commenced towards the east i.e. each one moving its way towards Swat, Lahore, Multan and Quetta respectively.

**2.4 Study Area**

The study area includes two geographically adjacent districts of Khyber Pakhtunkhwa (KP) province of Pakistan. i.e.Charsada and Mardan. Location map of the study area is provided inFig.2.

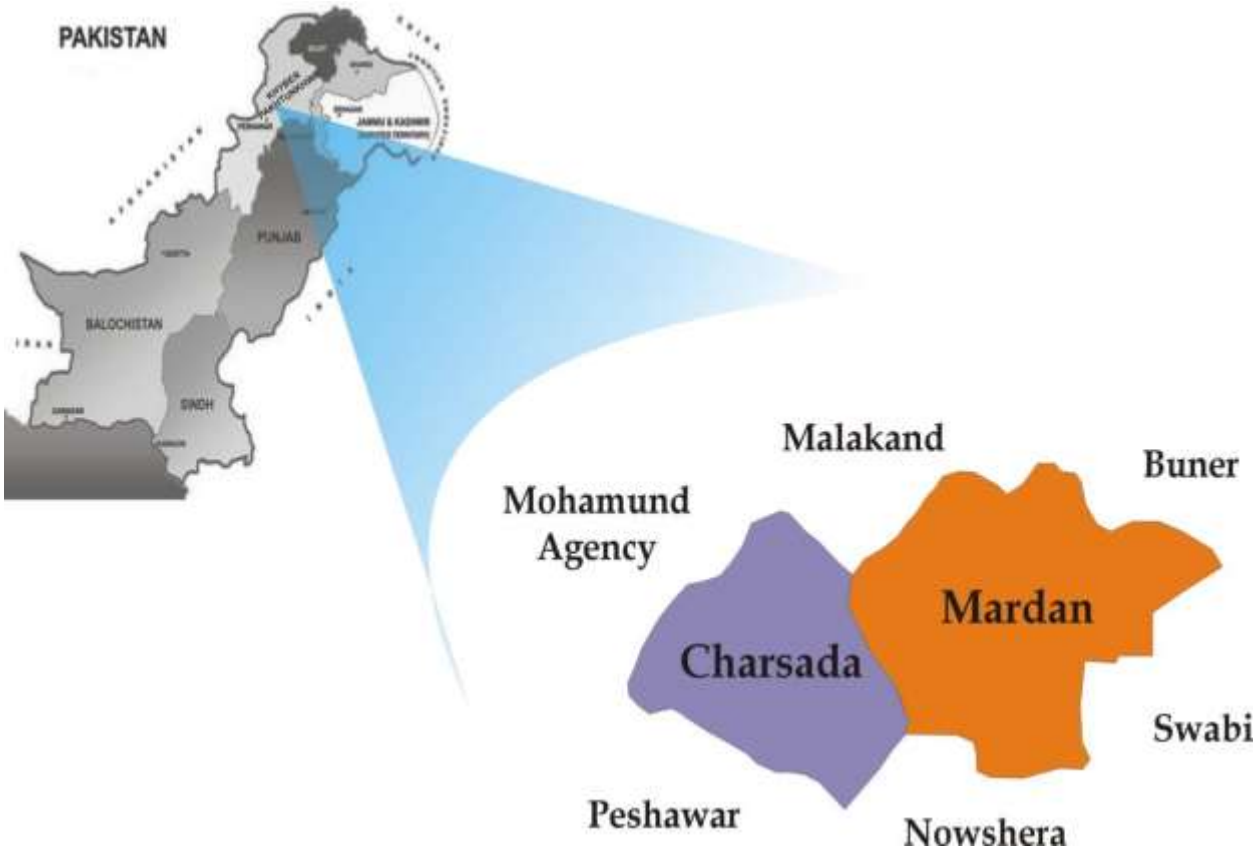



Fig.2. Location map of Charsada and Mardan district (Khyberpakhtunkhwa) Pakistan.

### 2.4.1 The Charsada District

Charsada is one of the 25 districts in the Khyber Pakhtunkhwa province of Pakistan. The town of Charsada was once part of the Peshawar region. Vast majority of the population of the area is Pathan.

#### 2.4.1.1 Geography

It occupies total area of 996 $km^2$ with total population of almost 10.2 million individuals according to 1998 census report, and 16.2 million according to 2014 report. It is situated at 34° 03" and 34° 38" north latitude and 71° 28" to 71° 53" east longitude. It is bordered by Malakand district at its north, Mardan district at its east, Nowshera, Peshawar at south and federally administrated tribal areas (FATA) and Mohmand agency at its west (fig.2). Being geographically adjacent to Mardan district, the climate of Charsada is similar to that of Mardan.

Administratively it is divided into three tehsils (Charsada, Shabqadar and Tangi) collectively consisting of a total of 46 Union Councils.

#### 2.4.1.2 History

It is considered as one of the most ancient and historical sites of Asia characterized by possessing several impressive mounds in the area. Charsada remained the part of the Gandhara kingdom that had been a part of $7^{th}$ strappy of Achaemenid Empire around

516 BC until it was overthrown by Alexander the Great in the 4th century BC. The city was detained by Alexander" s army in 327 BC. The Indian sovereign Chandragupta Maurya took over the control of Gandhara after the death of Alexander the great in 323 BC. After this, Maurya" s grandson Ashoka became an emperor and in his reign he built one stupa there of 2.5 miles perimeter which was later revealed by Hieun Tsang., a legendry Chinese Buddhist pilgrim in 630. After the conquest of Mahmud of Ghazni and arrival of Islam in 1026, the name Gandhara was faded away (Muhammadzai, 2012; Ali, 1993). **B**etween 250–125 BC, this area was also governed and ruled by the Bactrian Greeks which was later succeeded by the Indo-Greek Kingdom who ruled until 10 AD. Formerly, the town of Charsada was known as Pushkalavati (Lotus City). This name was first cited in the Hindu epic chronicle the "Ramayana". It was based on the cultivation of Lotus in this area that can be witnessed by archeological records and coins found from this area with the picture of goddess carrying a bunch of Lotus flower (Marshall and Vogel, 1903: 176). According to Ramanaya, Ramachandra" s brother Bharata founded two cities of Taksha and Pushkala on the names of his two sons, after the conquest of Indus valley (Dani, 1963). According to Vaidya, these two towns were named Taxila and Peukhlaoti in Greek history. This name was replaced by a Persian name "Hashtnagar" with the arrival of Muslims in this region in 11th century. Hasht meaning "eight", so its eight main villages Prang, Rajjar, Utmanzai, Umarzai, Tangi, Sherpao, Turangzai and CharsadaBazaar gave its name Hashtnagar. This name was first mentioned in Babur" s memories of 6th century (Babur, 1987: 55). Garrick was of the view that the name Hashtnagar was a combination of two words of two different languages

i.e. Persian and Sanskrit (Garrick, 1881-82). According to Cunningham, the name was derived from its original name "Hastinagra" that was named after the ruler of this area, Hasti or astes (Cunningham, 1963). Later on the name evolved into Charsada when Ilyas Khan Muhammadzai., an Afghan conquerorsettled here in 1109 and divided the land among his four sons. As in Persian the share of land was called Rasad, so it became Char-Rasad or four shares (Gopaldas, 1889). This area had been extended over a large area, and the entire vicinity is covered with vast ruins. Excavation was accomplished in district Charsada for about two months in 1902. Bala Hisar, an archeological site was exhumed twice by the Sir John Marshall in 1902 and by Sir Mortimer Wheeler in 1958.

According to Wheeler, "Bala Hisar was established by the Persians in the 6th century BC as a camp protecting the eastern edge of their territory" (Wheeler, 1962). Several interesting remains of coins and pottery ornaments were found. A series of post holes and ceramic remains and other archeological deposits recovered from Charsada excavations are dated to 1400 BC (Hussain, 1992). Subsequent epochs signify that other more permanent structures including stone-lined pits were also constructed at Charsada. Charsada people started developing an iron-working industry and started using ceramics between 14$^{th}$ to 6$^{th}$ century BC distinctive for this period in the gorge of Peshawar, Dir and Swat.

From 6$^{th}$ century BC to 2$^{nd}$ Century CE, Charsada remained the capital of Gandhara Kingdom being its administrative centre. This area was ruled by many foreign invaders including Persians, Mauryans, Greeks (Alexander), Greco-Bactrians, Scythians,

Parthians, Kushans, Huns, Guptas and Turks (Gazetteer, 1997-98).

**2.4.1.3 Tribes**

80% of the population of district Charsada belongs to Muhammadzai tribe which is sub-divided into groups locally called "khel".These sub-tribes(khels) have the status of khans of town. In addition several Sayyid families also exit. Mainly are Kakakhel Mian/Sayyid who claims to be the descendants of Sheikh Rehamkar Hazrat Kaka Sahib the most venerated Pir or Spiritual Figure among the entire Pathan population in Pakhtunkhwa Pakistan and Afghanistan. This tribe has gained special respect and honor among other Pathan tribes due to their virtuous nature and pious descent and ancestry. They are declared as land lords for having lands of prime importance in Charsada and neighboring districts.

**2.4.2 The MardanDistrict**

Pir Mardan Shah was a prominent religious figure of a small area which was named after him. Later a vast surrounding area merged into this small territory of Mardan.
This area was once a part of Gandhara kingdom being a fraction of Peshawar valley.

**2.4.2.1 Geography and Topography**

The district occupies total area of 1632 square kilometers at 34° 05‟ to 34° 32‟ north latitudes and 71" 48‟ to 72° 25‟ east longitudes. It is delimited by Swabi and Buner districts on east while Buner and Malakand protected area on north, Nowshera district on south and Charsada district and Malakand protected area on the west (fig.2). At

northern side the district is bounded by high mountains with highest points of Pajja or Sarka 2056m high and Garo or Pato at the height of 1816m above sea lea level, while there is a vast south western fertile plain with some low hills encumbered through it. There is a steep slope running through the Plain carrying rain water to the Kabul River from foot hills. Water bodies and streams flow from North to south direction. The important streams of the area are: Kalpani, Baghiari khawar, Muqam Khawar and Naranji khawar.Its climate is considerably hot and humid in summer with frequent dust storms at night. December and January are the coldest months with minimum temperature up to 0.5˚C. According to 1998 census, its total population was recorded almost 14.6 million with approximately 3.01% growth rate.

**2.4.2.2 History of Mardan district**

The area comprising Mardan district is a part of the Peshawar valley, which first emerged on map as part of the Gandhara kingdom. After invasion of Alexander the Great, the mists of murkiness began to disperse. The Alexander" s armies invaded the
Indus valley through two different routes, one via Khyber Pass and the other through Kunar, Swat, Bajaur and Buner in 326 B.C. After Alexander" s exodus, Chandragupta took over the control of this valley from 297 to 321 B.C. Buddhism was at its full swing in Peshawar valley for the period of the supremacy of the Buddhist emperor Asoka, Chandragupta" s grandson.In the reign of emperor Mehanda after the Greeks took over the valley, there was a revival of Brahmanism. After this, the Scythians and Indians followed and kept hold of the valley till the 7$^{th}$ century A.D.

By the end of 7th century, the Afghans invaded the valley. At that time Peshawar valley was governed by the rulers at Lahore. The Afghans united to the Gakkhars and over threw the Lahore rulers and took over the control of the hill country south of the Kabul and west of the Indus River. The area came under command of the Sultan Sabuktgin after defeating the Hindu ruler of Lahore Raja Jaipal in 10th century. After this the area had become a rallying point of several incursions and foray to interior India, for Sultan Mahmud Ghaznvi, the son of Sabuktgin. Ghaznavaid era came to an end when Pathans of Ghor took over in 15th century. The area was then invaded by Mughal emperor Babar via Khyber Pass in 1505 and remained under the authority of several Mughal emperors up to Aurangzeb. During his government the Pathan tribes revolted and in spite of continuous effort of two years from 1673-1675, to re-establish his authority Mughals were forced to agree with the terms set by Pathans for to retain their independence.All the territory west of the Indus including present day Mardan District was ceded to Nadir Shah by Mughals in 1738. Ranjit Singh after taking control of Attock in 1814 and Peshawar in 1818 handed over the command to Hari Singh Nalwa and moved further to Lahore. Sikhs ruled this valley until 1849 and were overpowered by the British Army in the Second Sikh War. Major Lawrence was appointed as the first Deputy Commissioner of Peshawar and Peshawar was declared an administrative district under the Government of Punjab. At that time the present Mardan district was merged with Peshawar district and later in 1909 Frontier Province was came into being. Peshawar district was branched off into Peshawar and Mardan districts in 1937

(Hastings, 1878).

#### 2.4.2.3 Ethnicity

Yousafzai Pathans constitute the majority of Mardan population. But in Lundkhwar valley of Tehsil Takhtbhai sizeable Mohmand population also exist. The origin of the Yousafzai tribe is traced back to two brothers Khakai and Gori. They gave their names to the two divisions of the tribes inhabiting the area near Kandhar. In the middle of the 13th century, the Khakal were consequently expelled by the second group of Ghoris and they settled in the area near Kabul. After getting enough wealth and fame and getting increase in population size, they divided into three clans, the Turklays, Gigyanis and Yousafzai. By the end of 15th century, the Yousafzai and Gigyani tribes started moving towards the plains of Peshawar and ultimately overthrew Dalazaks and dispersed into Buner district. They finally settled into Mardan district and the area was known as "Yousafzai Plain".

### 2.5 Tribes under Scientific investigation

#### 2.5.1 Muhammadzai

The Muhammadzai also pronounced as Mohammadzai, Mohammedzai, Mohmandzai, Mamanzai (Murray, 1899). The Sarbani Pathantribe should not be mystified with the Muhammadzai of the Barakzai Durrani lineage of Afghanistan. According to Pathan family history, the Muhammadzai Pathans claim to be descended from Qais Abdur Rashid who had three sons; Sarban, Bait and Ghourghusht. His son Sarban is known to be the ancestor of Muhammadzai of Charsada through his son Kharshban while Afghani Muhammadzai are said to be descendants of Sharkhbun,the Kharshbun's brother. One of the three sons of Kharshban named Zamand had a son named

Muhammad who is the actual founder of this tribe (Caroe, 1957; Rose, 1997).

Muhammadzai tribe is found primarily in Hashtnagar, an area in present day Charsada District. Initially it was said to have settled in Khorasan, but migrated to Pakistan and was given the Hashtnagar by theYousafzai tribe (Elphinstone, 1815). In 9th century the area near Qandhar named Pusheen was inhabited by the Zaman tribe. The area was later invaded by Tareen tribe and zaman tribe was expelled from their home land. Muhammadzai tribe before this expulsion started scattering from Afghanistan and reached Eastern province of Afghanistan Ningharhar. Where they had battle with Yousafzai tribe but defeated by them. Even after defeat they were not expelled from Yousafzai area of Ningharhar and were asked for help from Yousafzai, when they planned attack against Dalazaks. Muhammadzai joined Yousafzai but asked for award of territory of Hashtnagar which they achieved after their victory. They still retain their hold. Mir panda Khan, Malik Khyzer Khan and Hegi Khan were declared the tribe leaders. In 1526-1530, uniting with Mughal emperor Babur conquered Peshawar and overpowered Muhammadzai and Yousafzai. After Babur, Muhammadzai took over the control of the area. The geography of tribes is fundamental to their internal organization, because the subdivisions of the tribe and their respective villages are known by the same names.

### 2.5.2 Mohmand

For centuries between the basins of the Oxus River and Tarnak River in Ghazni, there lived a Mohmand tribe. This tribe has purely Afghan origin and follows strict pre Islamic honorific code "Pashtunwali". At the time of British invasion in Peshawar valley, the area comprising the present day Mohmand agency was under the influence of the local tribesmen and was also governed by Khan of Lalpura.

The Mohmand had many rebellions against the foreign invasions and specifically against British India. Many Khels of Mohmand tribe were given assurance in 1893 from the Emir of Afghanistan that they will not be endured severance because of their past association with Afghanistan, and they were known to be Assured Clans. Mohmand living in Charsada and Mardan district are said to be migrated from Mohmand area of Mohmand Agency, which is bounded by Charsada and the plains of Peshawar on East.
There are also several sub clans of Mohmand i.e. Khawezai, Tarakzai, Baizai and Halimzai which are further subdivided in several sub clans and divisions.

The famous Sanskrit grammarian and historian Panini also referred to the name of tribe the *Madhumants* (the modern day Mohmand) in his Ashtadhyayi (Grammar of Sanskrit) who colonized the northwestern areas, in the 5th century BC. (Roy *et al*. 2007).

### 2.5.3 Yousafzai

Yousafzai tribe is considered to be the largest Pathan tribe initially originated in

Qandhar, the present-day Afghanistan. Arrived and settled in Kabul in 1446 while migrating eastwards and there they remained under the influence of Turkic governor Ulugh Beg.

After further eastward movement towards Swat valley they faced rebellion with local Dilazak tribe. After a continuous effort of 20 years, under command of their chief Malik Ahmed Khan, the Yousafzai tribe along with some allied clans (Jadoon and Uthmankhel) ultimately succeeded to throw out Dalazaks from their home town and they were then pushed eastwards to the Hazara Mountains east of the Indus River, after the encounter of Katlang.

In 330 BC Alexander mentioned a tribe named "Isapzais". Similar name was alsopointed out by Baburin 16th century. Because of their non cooperative attitude, Akbar tried to subdue them under the command of Raja Bir Bar. Consequently Raja Bir Bar was killed by Yousafzai. It was due to their infuriated nature that they could not be completely brought under dominion of Mughal Empire until 1690 (Richards, 1993). Pir Babawas their first emir. After Akbar Shah's death in 1857, Akhund Ghaffur took over the control of state and ruled the stated until 20$^{th}$ century (Haroon, 2011).

The Yousafzai tribe is predominantly found in the districts of Swat, Shangla, Buner, Malakand, Mardan, Tor Ghar, Swabi, Lower Dir, Upper Dir in the Khyber Pakhtunkhwa province of Pakistan, besides some living in Battagram and Oghi (Mansehra).

They also inhabit areas other than Khyber Pakhtunkhwa in Pakistan e.g. a Brahui and Persian speaking Yousafzai clan has also been found in Mastung district Baluchistan. The predominant language of this tribe is Pakhto (Pashto), the northern variant of Pashto with the hard "kh" replacing "sh" in soft southern variant. Some Yousafzai lineages are found in Andhra Pradesh, Uttar Pradesh, Madhya Pradesh, Gaya, Bihar, Gujrat and Banglore (India).

**2.5.4 Kakakhel Mian**

The Kakakhel Mian is considered to be a prominent Sayyidclan. Their roots reach to Hazrat Ali bin Ismail bin Imame-Jafer Sadiq.Sayyid Kastir Gul an Islamic Sufi or wali is their probable ancestor. Another probable ancestor is Sheikh Rehamkar, a student of Sheikh Hazrat Akhund Adeen/Adyan Seljuki. The title of Rehamkar was given by Sayyid Abdul Wahab Akhund Panju Baba.

Kastir Gul was affectionately called "Kaka Sahib", so his descendants are entitled Kakakhel. The probable place of origin of this clan is a small village named "Kakasaib" in Nowshera District, KhyberPakhtunkhwaProvince from where they spread throughout the province especially in the nearby area of Charsada district (District census report, 1998).

The word "Kaka" means "uncle" and "Khel" means, "sons" or "children". So the word "Kakakhel" means the "Children of the Uncle". As Kakasaib was affectionately called "Kaka" by everyone in the village, so the name Kakakhel was given to his descendents.

Until late 13th century, Kakakhel Mian was least known clan because of their numerical scarcity beyond their borders. But they started getting prominence in 19th century by considerable increase in their population size. Their religious education and righteous nature contributed to high esteem among other Pathans. Kakakhel have been recognized as forefront fighters in battle against Sikhs (Official Gazette, 1870). During the British Raj, Kakakhel made a number of contributions to society. They proved to be highly competent civil contractors, soldiers, diplomats and police officers.

Kakakhel Mian has also made contribution to the independence movement and they have also served Pashto and Persian language for decades. Several writers and poets of Pashto literature belong to Kakakhel clan of Sayyid. The intellectuals and writers from Kaka sahib developed a union named "Milliah Rehamkaria" at the onset of 20th century where they donated a collection of about 4000 books, magazines and newspapers. Sayyid Bahadur Shah Zafar Kakakhel, a renowned intellectual compiled a dictionary comprising 45000 words. Members of this clan use "shah" as their surname (Gopaldas, 1874).

**2.6 Mitochondrial genome**

The length of Human mitochondrial DNA (mtDNA) is 16,569 base pairs that code for 13 proteins involved in oxidative phosphorylation, 2 rRNA and 22 tRNA molecules (Fig.3). Several other proteins i.e. RNA and DNA polymerases and RNA processing enzymes are of nuclear origin and are imported into mitochondria via double mitochondrial membrane (Cummins, 1998). Due to unsymmetrical distribution of Gs and Cs in both strands of the circular mtDNA, one strand is considered as Heavy (H) and other is Light

(L). Mitochondrial DNA is a circular molecule present in multiple copies per cell ranging from 100 to 1000. This attribute makes it useful in the presence of small quantity of sequence of interest, especially in case of fossil study. Because only an egg donates its mitochondria to the developing embryo with a few exceptions, in all humans mtDNA is maternally inherited (Giles *et al*. 1980; Cummins, 2000). The mitochondron of sperm is degenerated by the enzymatic activity of ubiquitin (Thompson *et al.,* 2003). As the individual harbors only one type of multiple copies mitochondrial DNA, said to be homoplastic but become Heteroplasmic when mutation occurs in any of the mtDNA copy during segregation (Cavalier *et al.,* 2000; Alonso *et al.,*

2002). Unlike nuclear genome there is no recombination mechanism in mtDNA(Meriwether*et al*. 1991). Similarly mtDNA is also lacking of repair mechanism that consequently leads to the much higher mutation rate than nuclear genome sequences (Brown*et al*. 1979, 1982; Saccone2000). This attribute is the basis of sequence variation among individuals (Olivio *et al*., 1983). Its high copy number renders it useful tool in DNA profiling for forensic investigations (Allen *et al.,* 1998).

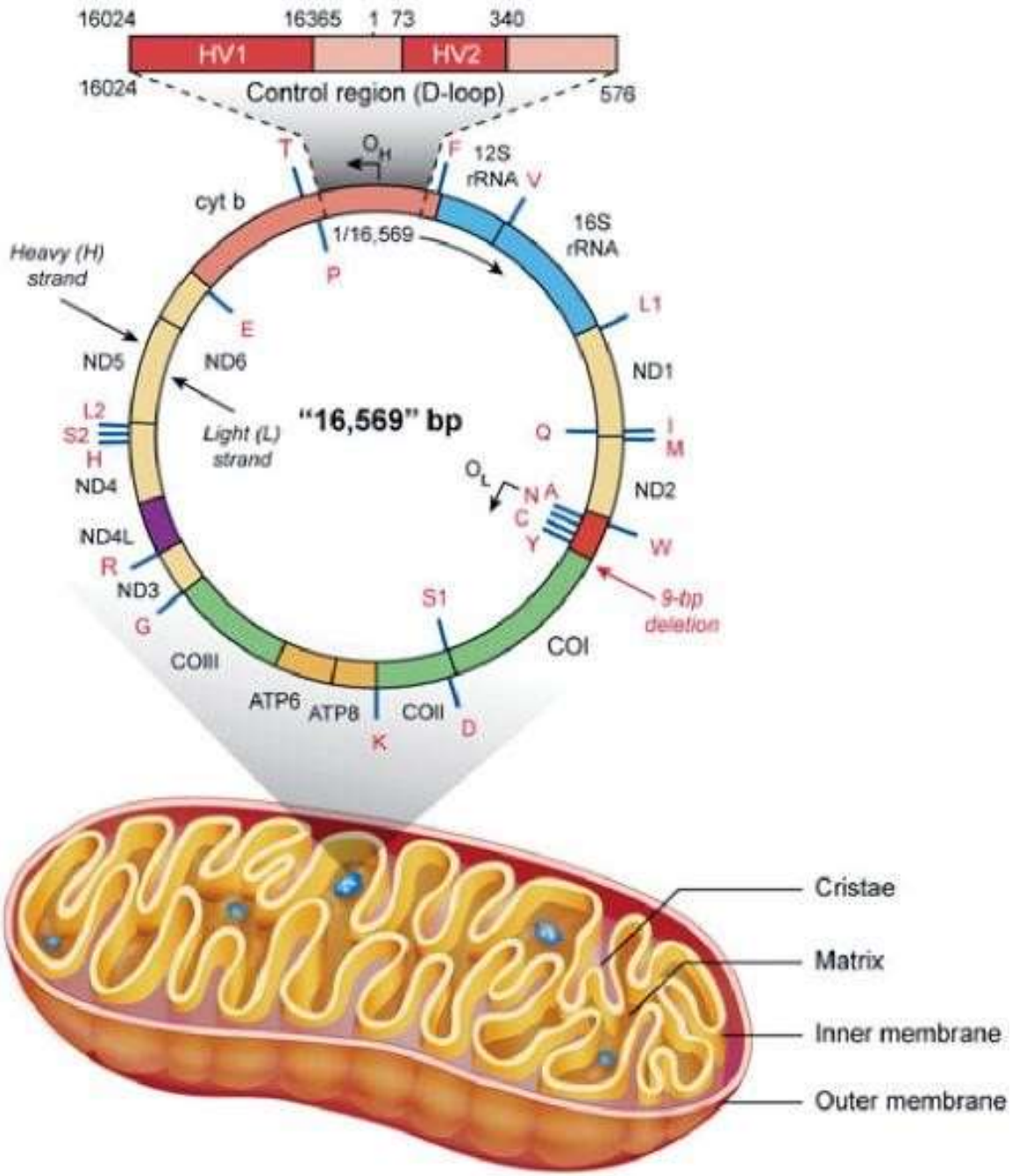



Fig.3.The16569 bp, circular human mitochondrial genome. The control region (D-loop) with the two hyper variable regions: HVI and HVII which are commonly used in Population genetic investigations (Lembering, M. 20013).

### 2.6.1 Tracing human migration from mtDNA

A few decades back, the only possible way of tracing human migration and to get insight into the evolution was to expose and uncover the skeletal remains. However, several breakthroughs have come up in the last two decades for the anthropologists with the advancement in the field of molecular biology. Until late 19th century, analysis of biomolecules like proteins was proved to be the useful tool for evolutionary studies but with less accuracy. With the discovery and introduction new genetic tool and molecular techniques, to get deep insight into the intra and inter population variation has been easier for the anthropologists to formulate new evolutionary and migration theories. DNA whether nuclear or extra nuclear has a tendency to record set of mutations in its sequences taking place with the passage of time so, act as a biological calendar. The role of mitochondria in several biological and physiological processes has long been known but their role in tracking evolutionary history has recently been studied in the last three decades. The rate of sudden changes/mutations accumulated over time, defines the rate of evolution that can proves to be a useful tool to elucidate the time and age of divergence. e. g. Genetic and paleontological data suggested the common lineage of humans and Chimpanzee and their time of divergence was nearly 5million years ago (Ingman *et al*., 2000). Whole mitochondrial genome doesn" t evolve simultaneously at same rate (Pesole *et al*., 1999). The highest rate of mutation is in non coding triple stranded stretch of mtDNA called D-loop or displacement loop (Wallace *et al*., 1995) and said to have three hyper variable regions, HVI, HVII and HVIII (Stoneking, 2000). This region comprises about 7% of mtDNA and has been studied and scrutinized in detail (Ingman *et al*., 2000).

The rate of mutation in D-loop region is 10 times higher than in coding region (Howell *et al.,* 2007). On average 1-3 bp per 100bp varies among two unrelated individuals at random in their D-loop sequences (Piercy *et al.,* 1993). Although there are some kinds of mutation which make comparisons difficult e.g. back mutation and parallel mutation. Similarly the presence of hot spots for mutations also renders the comparison difficult but sequencing of complete human mitochondrial genome has offered many polymorphic sites for comparison. The HVI region ranges from 16024-16365 np and HVII ranges from 73-340 np while HVIII ranges from 438- 574 np. To annotate the population differences, mtDNA sequences are compared with (rCRS) Revised Cambridge Reference Sequence (Andrew *et al.,* 1999). It was discovered on the basis of Analysis of HVR1 and HVR2 of mtDNA that, around 500,000 years ago Human and Neanderthal split took place (Krings *et al*., 1997) and it was also hypothesized that there is no contribution of Neanderthal in modern human mtDNA gene pool (Krings *et al*., 1997, 1999; Ovchinnikov *et al*., 2000). The first phylogenetic tree for mtDNA based on Restriction Site Polymorphism was constructed by Cann *et al* in 1987 and the maternal most recent common ancestorwas called

Mitochondrial Eve originated around 200,000 years ago in Africa (Cann *et al.,* 1987). Like Y-STRs, being haploid markers mtDNA polymorphisms are best candidate for phylogenetic and population origin investigations including tracing human migrations. Such polymorphisms are geographically dependant as the mutations accumulated in mtDNA lineages during the course of time (Ingman *et al.,* 2001). These maternally inherited variants constitute haplotypes. A specific set of SNPs defines a haplogroup and

various haplotypes belong to same haplogroup on the basis of shared haplogroup defining SNPs (Torroni *et al.,* 1996). Haplogroup L0 - L7 are considered to be the most ancient haplogroups of African origin (Quintana *et al.,* 1999) while the haplogroup L3 is considered to be the first common ancestor of M and N haplogroup which later evolved into many variants during and after out of Africa human migration (Richards *et al.,* 2006; Chen *et al.,* 1999) A generalized sketch of human migration as reflected in mtDNA is provided in Fig.4.

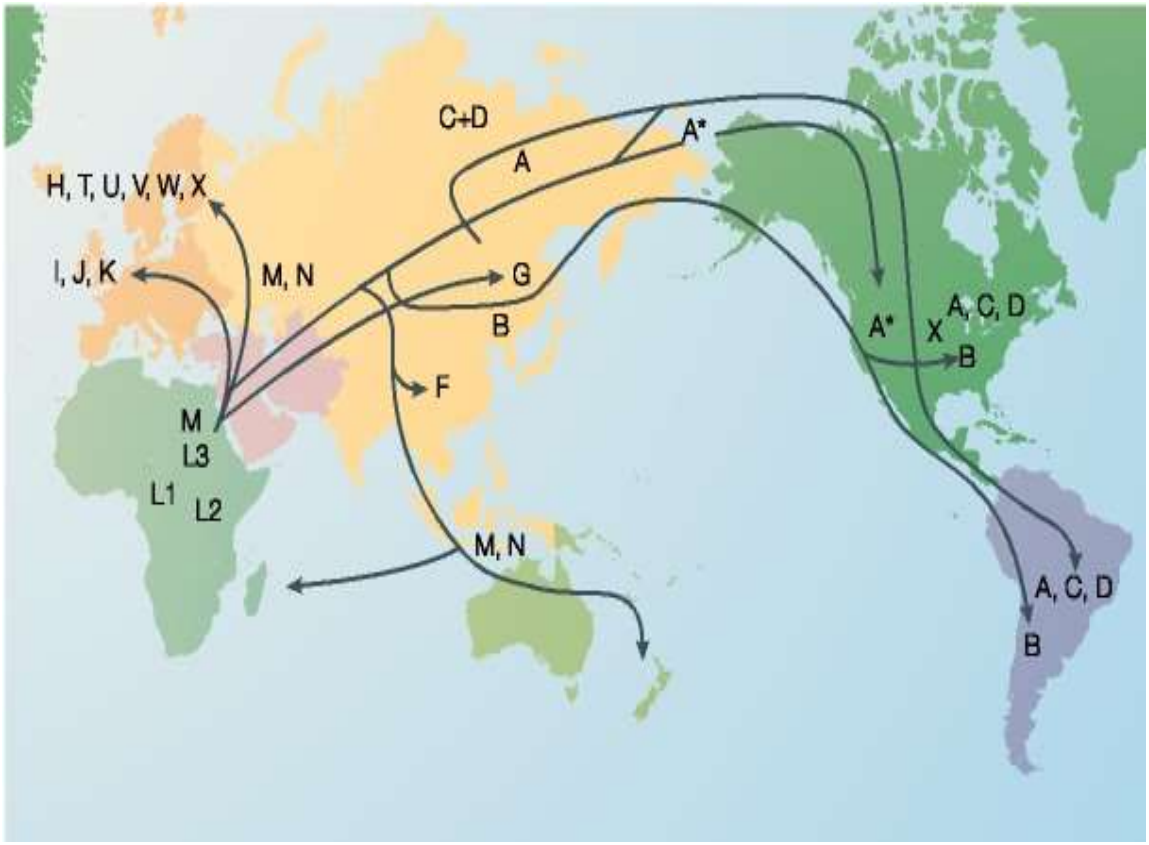



Fig.4. An Overview of worldwide distribution of mtDNA haplogroups (Schriver & Kittles, 2004).

### 2.6.2 mtDNA Haplogroups in South Asia

#### 2.6.2.1 Haplogroup M

Macrohaplogroup M, along with its sibling N, account for all non-African mtDNAs. This haplogroup originated during the human migration from Africa 57-75Kya (Chandrasekar *et al*., 2009). An age estimate of 49,400Ky (95% CI; 39,000-62,200 years) is provided by Soares *et al*. (2009) in south Asia and in East Asia 60,600Ky (95% CI; 47,30074,300 years). Other approximations include 55-73Ky for the lineage among African populations (Chen *et al*., 1995) and 69.3 ±5.4Ky among Chinese populations (Kong *et al*., 2003). Due to its significant contribution and distribution within the Indian population i.e. 60-70%, the haplogroup M is considered to be a south Asian lineage (Chandrasekar *et al*., 2009; Disotell, 1999), but lower frequencies of this haplogroup has been estimated in Central Asia and western Eurasia. It is believed to have arrived in the Indian subcontinent via the Southern Route migration from Africa (Disotell, 1999; Macaulay *et al*., 2005; Torroni *et al*., 2006; Chandrasekar *et al*., 2009; Kumar *et al*., 2009). Sub-group M7 is a common lineage found in East Asian populations such as Korean-Chinese, the Han (Beijing) of China, Mongolians, Koreans (Jin *et al*., 2009) and Japanese (Asari *et al*., 2007). The Hazara, Balouch and Pathan populations of Afghanistan exhibit 15%, 13.3% and 7.1% frequencies respectively of this haplogroup (Whale *et al*., 2012). within the Indian subcontinent its Frequencies have been found ranging from 26%-64% among different caste and tribal populations (Kivisild *et al*., 1999; Quintana-Murci *et al*., 2004) while the frequencies found within the Afghan populations seem to resemble

frequencies found elsewhere in Central Asia (Whale *et al.*, 2012).

#### 2.6.2.2 Haplogroup N

The second macrohaplogroup to have diverged from the African lineage L3 is haplogroup N. Its age has been estimated at 64.6 ±6.8Ky (Kong *et al*., 2003), 61,900 YBP in west Eurasia (95% CI; 49,200-75,000 years), 71,200YBP in South Asia (95% CI; 55,80087,100 years) and in East Asia, 58,200YBP (95% CI; 44,100-72,800) (Soares *et al.*, 2009). It is considered to be the ancestor of many haplogroups found in Europe, Middle East, Asia and the Americas (among the Amerindians). The origin of this lineage occurred soon after or probably during the migration out of Africa, and is typically considered a southwest Eurasian lineage (Kivisild *et al*., 1999; Quintana-Murci *et al*., 2004; Nasidze *et al*., 2006, 2007). This haplogroup is comparatively common in western Eurasia and is also present in Europe. A frequency of 5.3% has been reported in eastern Crete (Martinez *et al*., 2008) while the combination of haplogroups N, I, W and X constitute approximately 9% of the Finnish population (Hedman *et al*., 2007). This haplogroup is also found in the Near East and northeast Africa nearly ~13% in Egypt, ~10% in Israel,
Syria and Jordan, ~5% in Iraq and ~23-44% among Iranian populations (Nasidze *et al*., 2008). The Hazara population of Afghanistan exhibits a frequency of 7.5% and the Tajiks 10.5% of this haplogroup (Whale *et al*., 2012). Elsewhere in Central Asia and western Eurasia, the frequency of N haplogroup ranges from 2.3% in the Tajiks of
Tajikistan (Derenko *et al*., 2007) to 20% in the South Caspian region in Iran (Comas *et al*., 2004). The greater frequencies of this haplogroup appear to occur in the more western

populations rather than in Central Asian or South Asian. Haplogroup M is affluent in South Asia, however haplogroup N appears to be lacking from the mtDNA landscape with frequencies of 2.6% and 2.9% in the Brahui of Baluchistan and Gujarati of northwestern India, and 3% in both Pakistani and Makrani populations and 7.7% within the Han Chinese population (Yang *et al*., 2011). The frequencies exhibited in Central Asia are similar to those found in the Hazara and Tajiks with Uzbeks harboring 7.1% (Quintana-Murci *et al*., 2004) and the Turkmen population 10% (Comas *et al*., 2004).

#### 2.6.2.3 Haplogroup R

As a descendant of the macrohaplogroup N, the haplogroup R also diverged soon after the modern human migration out of Africa. Along with the macrohaplogroups M and N, R is one of the founder lineages for Eurasian settlement ~60-65Kya (Torroni *et al*., 2006). It has an estimated age of 59,100 years in west Eurasia (CI; 47,100-74,100 YBP), 66,600 years in South Asia (CI; 52,600-81,000 YBP) and 54,300 years in East Asia (CI; 41,200-67,800 YBP) (Soares *et al*., 2009), and 62.3 ± 6.3Ky (Kong *et al*., 2003). This haplogroup is a typical west Eurasian and South Asian lineage primarily due to its early divergence from haplogroup N in this region and can be characterized by a MboII site gain at np 12704 caused by a transition at np 12705. , The haplogroup R makes up for almost less than 3% of the maternal gene pool of Finnish population (Hedman *et al*., 2007). It has often been recorded in Central Asia, South Asia and western Eurasia; however its distribution is not uniform throughout these regions (Whales *et al*., 2012).

The Karakalpak population presents a frequency of 10% (Comas *et al*., 2004), while in the Gujarati population of northwest India it appears in 8.8% of mtDNAs and in 1.8% of Georgians (Quintana-Murci *et al*., 2004). Within the Afghan populations, the Pathans exhibit 28.6%, Tajiks 15.8% and the Hazara 7.5% of this haplogroup. However, the greatest frequency was found in the Uzbeks at a frequency of 20%, while in the south Caspian region, haplogroup X was found in 2.4% of Persians and 9.5% of Mazandrians. The presence of these three lineages within the Afghani populations and the adjacent populations from Iran, Central Asia and the Indian Subcontinent may be attributable to this region being the initial territory where haplogroups M, N and R settled following the human dispersal from Africa. Despite each lineage sharing similar coalescent ages, haplogroup M is prominent among South Asian populations, particular among those in southern India in Andhra Pradesh where its frequency has been recorded at 64% (Whales *et al*., 2012). While the overall frequency of 61% for this haplogroup was recorded in Pathan Ethnic group of Pakistan (Rakha *et al*., 2011).

An updated mtDNA haplogroup phylogenetic tree(Fig.5) is being used for population haplogrouping and classification on the basis of origin (Van *et al.,* 2009).

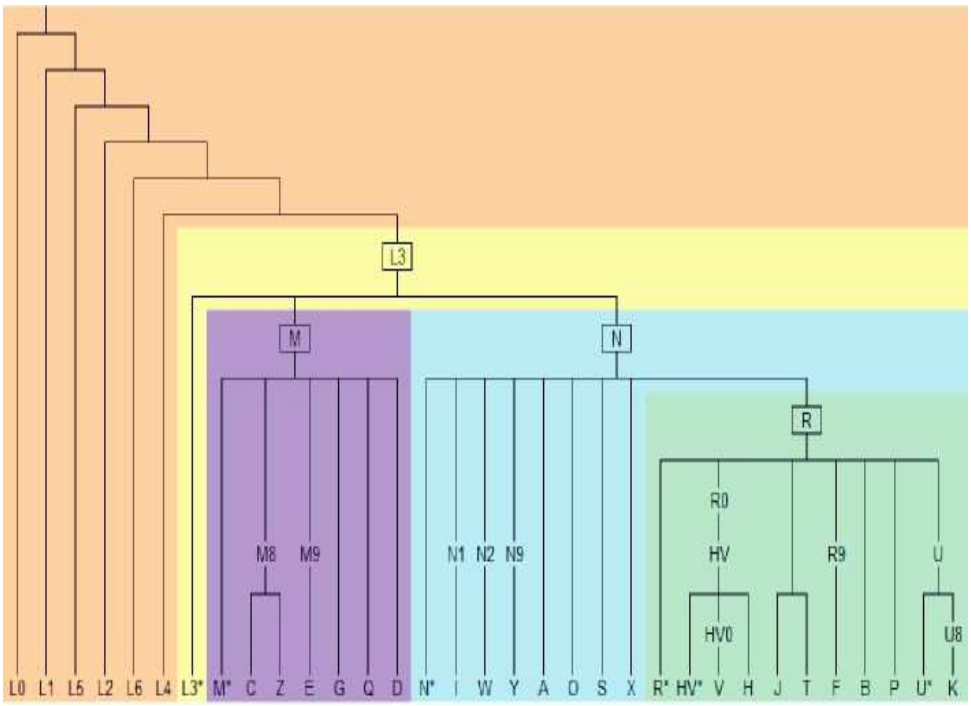



Fig.5. An Overview of the mtDNA haplogroups (Van *et al* 2009).

**2.7 The Y-Chromosome**

The human Y-chromosome is being considered as an evolutionary artifact of the X chromosome. Besides its ability to define gender, it has some other practical significance by carrying many functional genes essential for normal male development. The X and Y-chromosomes are considered to be the true homologues roughly around 300 million years ago (Lahn *et al.,* 2001) after successive mutations especially deletions reduced the Y chromosome to the size of 50 Mega bases, but sequence homology between to chromosomes still persist up to some extent.

### 2.7.1 The Structure of the Y-Chromosome

The smallest among all the chromosomes of human genome is the Y-Chromosome. Its average size is roughly around 60Mb. 24Mb of which comprises the euchromatin region and around 30Mb comprises the heterochromatin region, collectively they are referred to as the MSY (male specific region) or NRY (non-recombining region) constituting about 95% of the Y chromosome (Butler, 2003).Its large number of short tandem repeats (STRs) and single nucleotide polymorphisms (SNPs) contribute to the formation of a powerful and informative haplotyping system that has long been used in human identification (Jobling *et al.,* 2001). A sketch of the Y-chromosome is given in Fig. 6.

Several genealogical and evolutionary studies besides Forensic DNA typing have made use of continuously growing reliable Y- Short Tandem Repeats (STRs) and Single Nucleotide Polymorphism (SNPs) markers (Butler, 2003). Almost 99.99% of the Y-chromosome is inherited as a single unit from father to the son in paternal fashion (Butler, 2003). The entire chromosome remains intact and all its markers along its entire length remain strongly linked to each other. In this way a single chromosome acts as a single haplotype that can be used to trace the human migration and evolutionary events (Butler, 2003). The majority of Y chromosome exhibits lack of recombination and as a result of this attribute the haplotypes passes intact generation through generation and helps to preserve a simple historical record of the descent and phylogeny (Helena *et al.,* 2007). Unlike rest of the chromosomes mutation is the only source of variability in this case.

Due to small effective population size, i.e. one quarter to that of autosomes and one third to the X-chromosome, comparatively low sequence diversity is also expected on Y-chromosome than on any other (Hammer, 1995; Thomson *et al.*, 2000).

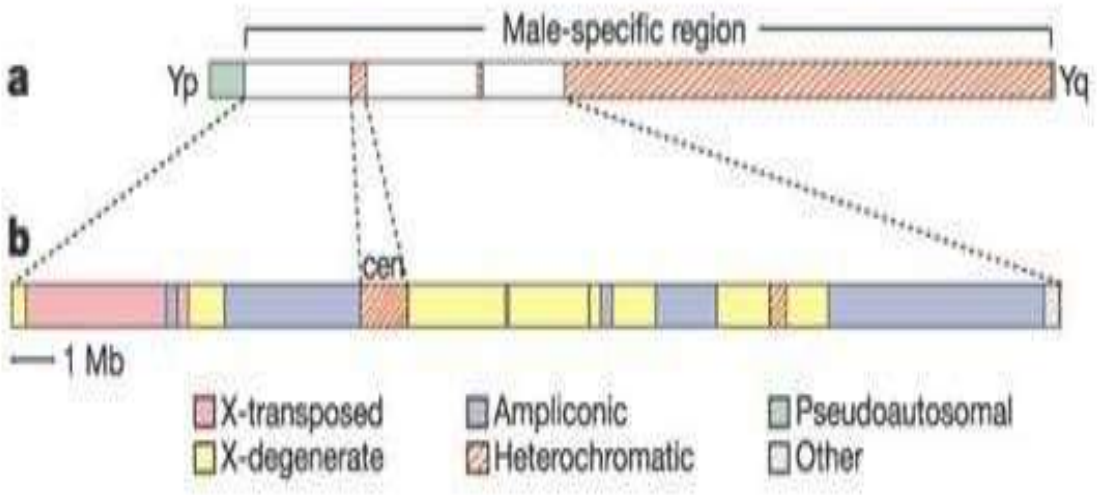



Fig.6. Structure of the various regions of the human Y-chromosome (Skaletsky *et al* 2003).

**2.7.1.1 Pseudoautosomal regions (PARs)**

The telomeric proximities positioned at the distal part of the short arm (Yp) and long arm (Yq)are called Pseudoautosomal region 1 (PAR1) and Pseudoautosomal region 2 (PAR2) respectively. The approximate length of PAR1 is 2.5Mb while PAR2 is less than 1Mb in length (Skaletsky *et al.*,2003). The Y chromosome recombine with its counterpart the X chromosome only at these regions during meiotic division in a similar fashion to that of the autosomal loci (Butler, 2003).

### 2.7.1.2 Heterochromatic region

The 30 Mb regions on (Yq) long arm of Y chromosome comprising two repeat sequences i.e. DYZ1, DYZ2, alphoid sequences and several satellite sequences, is called heterochromatic region on Y chromosome (Skaletsky *et al.*, 2003; Gusmao *et al.*, 1999). These sequences are clustered tandemly near centromere and plays important role in spermatogenesis (Gusmao *et al*., 1999). In addition to this, the Y-chromosome also carries two DYZ1 repeat fragments each 2.1 Kb in length and two other fragments i.e. the Y-specific (YS) and the non-Y-specific (NSY) (Gusmão *et al.,* 1999) each 3.4Kb in length. These fragments consist of a single array of pentameric satellite sequences with the repeat sequence of $^{5"}$ TTCCA$^{3"}$ in 800 to 4000 copies (Gusmão *et al.,* 1999).

### 2.7.1.3 Male specific region (MSY) or non-recombining region (NRY) of Y- chromosome

Almost 95% of the Y chromosome is responsible for sex determination in males called NRY (Non recombining region on Y chromosome) or MSY (Male specific region on Y chromosome) region. It consists of a mosaic arrangement of Heterochromatic and Euchromatic regions (Skaletsky *et al.,* 2003). The euchromatic region is further divided into three categories, i.e. the X-Degenerative Region, the X-Transposed Region, and the Ampliconic Region. Ampliconic region carries the total of 156 transcription units (Skaletsky *et al*., 2003).

#### 2.7.1.4 X-Degenerative Region of Y chromosome

The 20% of the euchromatin consist of eight sequence blocks on Yp and Yq comprising almost 8.6Mb of sequences, called X-degenerative region on MSY. Almost 27 X-linked single copy genes are located on this region. Out of which 13 genes are nonfunctional Pseudo genes while remaining 14 are Y linked functional genes. Some of which are homologous to X linked genes (Skaletsky *et al.,* 2003).

#### 2.7.1.5 X-Transposed Region of Y chromosome

X-transposed region of 3.4 Mb on the short arm of Y chromosome (Yp) comprises almost 15% of the MSY euchromatin and occurs in two sequence blocks (Skaletsky *et al.,* 2003). 99% of the X-transposed region sequences are homologous to the X chromosome long arm (Yq) sequences. This region possesses lowest gene density and highest interspersed repeat density and it doesn" t recombine unlike PARs (Skaletsky *et al.,* 2003).

#### 2.7.1.6 Ampliconic Region on Y chromosome

30% of the total MSY euchromatin constitutes the Ampliconic region of 10.2Mb, which consists of seven sequence hunks on both the short arm (Yp) and the long arm (Yq). These sequences possess the highest gene densities of the three classes of the MSY euchromatin and the lowest gene density of LINE1 and interspersed repeat elements
(Skaletsky *et al.,* 2003). Nine protein coding gene families on MSY have been identified.

Among them, XKRY, PRY, VCY and HSFY have two copies, BPY2 exists in three copies while DAZ and CDY exist in four copies and six copies exist of RBMY gene. All of them are involved in male fertility.

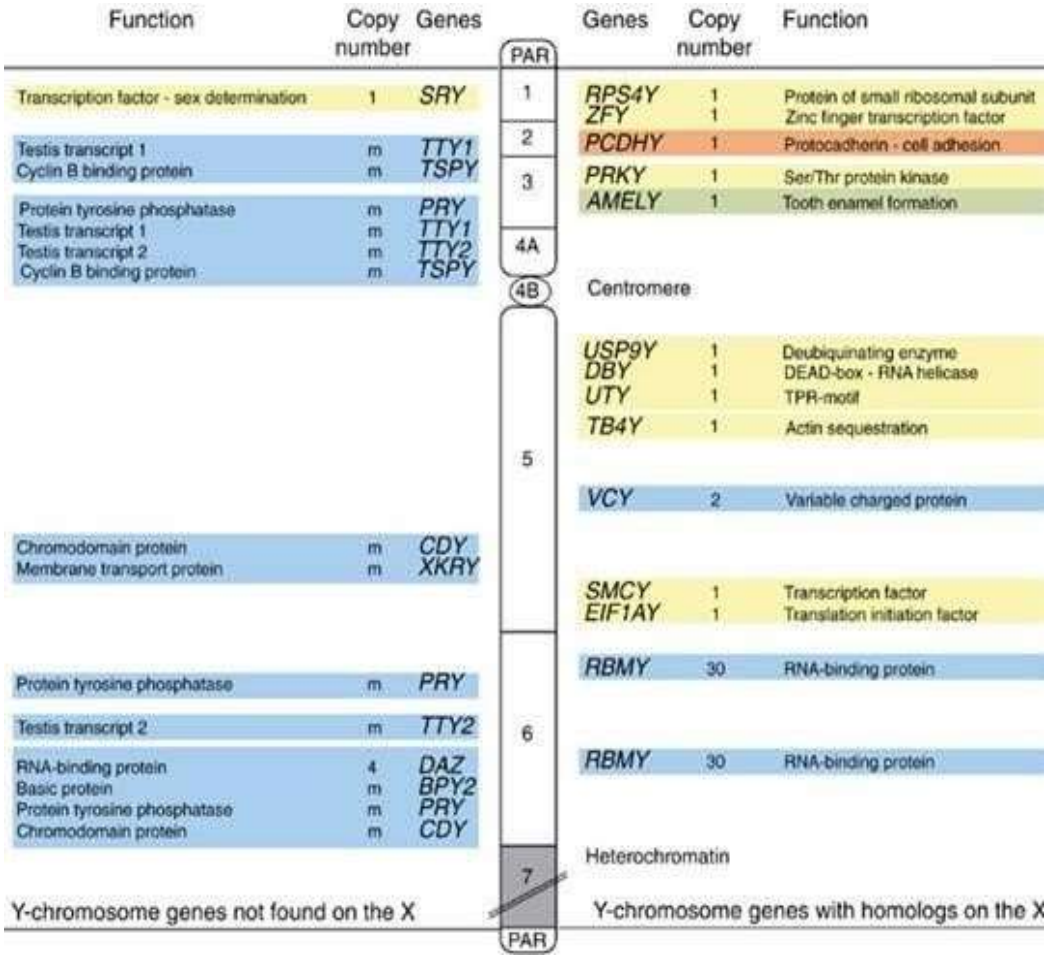



Fig.7. Protein coding genes on the Male specific region on Y chromosome with homologues on the X-chromosome (right side) and genes not found on the Xchromosome (left side) (Skaletsky *et al.*, 2003).

### 2.7.1.7 Palendromic DNA sequences

There are eight long palendromic sequences in ampliconic region on Yq, denoted by P1-

P8 ranging in length from 9Kb to 1.45KB. These palindromes are separated by 2Kb to 170Kb spacer sequences (Rozen *et al.,* 2003). P1 is the largest palindrome of 2.9Mb. It consists of two secondary palindromes P1.1 and P1.2 within it, each 24Kb in length. The total length of all the eight palindromes is approximately 5.7Mb that is about 25% of the total MSY euchromatin (Skaletsky *et al.,* 2003).

#### 2.7.1.8 Sex determining region on Y chromosome (SRY)

In the euchromatic region of the MSY, there is a region adjacent to the pseudoautosomal region on the short arm of Y chromosome called the sex determining region (SRY). It actually controls the male sexual development by encoding for (TDF) the testis determining factor (Sinclair *et al.,* 1990).

#### 2.7.1.9 Amelogenin gene (AMEL)

The Amelogenin gene (AMELY) on the Y-chromosome and its homologue (AMELX) gene on the X-chromosome, plays important role in tooth bud growth by enamel production (amelogenisis) and dentine development (dentinogenesis) (Butler, 2005). This locus is primarily used in gender differentiation of male/female mixed samples which is not only a practical tool but of requirement of extreme importance in forensic caseworks particularly in cases of sexual assault, skeletal remains and old blood blots which require swift, sensitive, reliable and accurate means of exploration and analysis (Butler, 2005).

### 2.7.2 Polymorphisms on the Y-chromosome

In order to trace the human migration, evolutionary history and population admixture, the paternally inherited Non recombining region on Y chromosome (NRY) has extensively been used since last two decades. This region on Y chromosome holds various polymorphisms with varying mutation rates that could be used as apowerful tool for scientists to achieve their research goals. The genetic markers on NRY region of the Y-chromosome are physically linked and are transmitted to the males of next generation i.e. in patrilineal fashion. This linkage and lack of recombinationresults in the lack of independent assortment of these markers and increased susceptibility to genetic drift due to small effective population size as compared to autosomes. Drift enhance the differentiation and discrimination between Y-chromosomes in different populations and is useful to investigate the past events in a specific population. But genetic drift influences the haplotype frequency in a population over time. Clusteringpattern is also influenced by the behavior of males bearing Y-chromosome, e.g. Patrilocality (Murdock, 1967; Burton, 1996). Patrilocality is mostly practiced in South Asia (Kayser *et al.,* 2001).

#### 2.7.2.1 SNPs the Bi-allelic Markers

The most abundantly occurring markers are the Bi-allelic markers which include the ALU insertion (YAP-DYS287) and considerable number of SNPs (single nucleotide polymorphisms) (Underhill *et al.,* 1996). SNPs are SNPs are the single nucleotide change

in the base-pair sequence as consequence of substitution, transversion or transition. More than 33,000 Y-SNPs have been discovered and used so far for population phylogenetic studies and forensic investigations (Hallast *et al.,* 2015). The most potential use of SNPs is the construction of haplogroups in population genetics and evolutionary studies. The most important is the appropriate and careful selection of SNPs, as different SNPs can define the same haplogroup (Sanchez *et al*., 2004).

#### 2.7.2.2 Alu Polymorphisms on Y chromosome (YAP)

Allele variants are created by mutations as a consequence of insertions or deletions of usually one nucleotide at specific loci along the Y-chromosome. One such important insertion is Alu polymorphism on Y chromosome (YAP) consisting of approximately 300bp (Hammer *et al.,* 1994).

#### 2.7.2.3 Multi-allelic markers (Micro and minisatellites)

Microsatellites are sequences with repeat units of 2 to 7 BP in length and mostly they are short tandem repeats (STRs) (Ellegren, 2000; Imad *et al*., 2014). The type of DNA sequence is defined by the length of core repeat and number of repeat units (Butler, 2012). They may be di, tri, tetra, Penta or Hexa nucleotide repeats. Penta and Hexa nucleotide repeats are considered as microsatellites. STRs have potential to be used to study evolution and migration foot prints as well as for resolution of medico legal cases (Walkinshaw *et al*., 1996).

Several hundreds of microsatellites or Y-STRs of potential importance in forensic caseworks and evolutionary studies have been identified on the Y-chromosome (Gil *et al.,* 2012). Number of repeats is highly variable among individuals so are best candidates for population genetic studies (Budowle, 1995; Butler *et al.*, 2009; Mohammed and Imad, 2013).

Similarly minisatellites are also present on Y chromosome that varies in length from 10 - 60 base pairs (Gusmão *et al.,* 2005). Number of repeats defines the allelenomenclature at specific locus. Changes occur occasionally in the number of repeats of these STRs or minisatellites. The most probable cause of intra allelic mutations in microsatellite is the replication slippage (Jeffreys *et al.,* 1998). Other than mutation, selection is another source of haplotype diversity in population. Paternal lineage can be traced by tracking these changes (Gusmão*et al.,* 2005). The rate of mutation is high in minisatellites loci, approximately equals to 6-11% per generation, while in microsatellites it is nearly equals to ~0.2% per generation for Y-STRs. On the other hand the mutation rate for SNPs is significantly low (Dupuy *et al.,* 2004). Multi-allelic markers have proved to be extremely useful in differentiating Y-chromosome haplotypes resulting in fairly high resolutions (Butler, 2003). Major fraction of the Y-STR loci is located on the long arm (Yq) of the Y-chromosome. Almost 25.3% of STRs are at Yq11.221, 16.6% of the STRs are at Yq11.222 and almost 18.4% of Y STRS are positioned at Yq11.223 (Hanson*et al.,* 2006).22.1% are found on the short arm of the Y-chromosome at Yp11.2 and three loci

DYS716, DYS707 and DYS631 are located in the centromeric segment (Hanson *et al.,* 2006). There are several advantages of using Y-STR markers over autosomal markers with a few limitations.

#### 2.7.2.4 Unique and recurrent event polymorphisms

Y chromosome markers are classified on the basis of their mutation rates. e.g. recurrent and unique event polymorphism (Hurles and Jobling, 2001). SNPs, YAP and indels exhibit low mutation rates while microsatellites and minisatellites exhibit high mutation rate and generally called "unique event polymorphisms" (UEPs). These are rarely occurring polymorphisms that are assumed to be population specific and occurs at specific loci as a consequence of evolution (Hurles and Jobling, 2001). Binary polymorphisms of distinctive origin (haplogroups) are pooled into monophyletic composite haplotypes and are phylogenetically branched off from the single parsimonious tree. Hurles and Jobling first time estimated the human ape divergence time on the genetic evolutionary clock. And the TMRCA (Time to most recent common ancestor) of existing human Y chromosome has been estimated to be between 50,000 and 200,000 years ago (Hurles and Jobling, 2001).

### 2.7.3 Applications of Y-chromosomal polymorphism

Several hundreds of the Y-STRs have been discovered so far are available as a traditional tool for forensic and population genetic investigations but only those with high genetic diversity and variance are proved to be more useful in such cases (Hanson *et al.,* 2006).

Rapidly mutating Y-STRs have ability to differentiate and discriminate between brothers in more than 60% of cases (Ballantyne *et al*., 2012).

#### 2.7.3.1 Y polymorphism, a tool in Forensic investigations

Y-chromosomal polymorphisms, primarily the STRs have been a useful discerning tool in forensic studies specifically in cases of sexual assault. They have been playing a substantial role in generation of valid genetic profiles acquired from male suspects of gang rapes, male /female or male/male sexual assault cases (Hanson *et al.,* 2006).

#### 2.7.3.2 Paternity testing casework

The uniparental inheritance along the patrilines makes the Y chromosome and Y–STRs useful for tracing the paternity of the male offspring. For this purpose the Y –DNA from any male relative in the lineage can be helpful (Gusmao *et al.,* 1999). A key feature in forensic investigation and paternity testing is the accurate construal of genetic profiles considering the probability % rates of possible STR mutations that could influence the inclusion or exclusion of paternity of an alleged father (Kayser *et al.,* 2001). Different sets of Y-STR markers are being used for this purpose. Y-STR haplotypes characterize information from the non-coding lineage of the Y-chromosome shared among several males along the paternal line. They do not provide individualization unlike that of autosomal STR loci and proved to be more valuable and conclusive in paternity tests of male subjects where autosomal and other STRs couldn" t work (Gusmão *et al.,* 2005).

#### 2.7.3.3 Implication of genealogical and Evolutionary investigations

Several sets of minisatellites, microsatellites or STRs and biallelic markers are the best candidates for evolutionary and genealogical investigations due to their typing simplicity and high level of diversity. Network analysis of the haplotypes based on these markers has proved to be more informative for evolutionary and phylogenetic studies (Gusmão *et al.,* 1999). Several hundreds of SNPs and Y-STRs are being used to investigate the biogeographical ancestry of different populations and sub-populations (Underhill *et al.,* 2001; Ali *et al.,* 2003).

Due to very low mutation rates and lack of recombination, SNPs are in strong agreement for the construction of maximum parsimony tree (Jobling *et al.,* 2001).

### 2.7.4 Y-Haplogroups

Human Y-chromosome haplogroups are defined by Y-chromosomal polymorphism specifically by SNPs. Each haplogroup represents the branch of Y-chromosome phylogenetic tree. The patrilineal most recent common ancestor of all humans is considered to be the "The Y-chromosomal Adam".

A system of naming Y DNA haplogroups had been developed by (YCC) YChromosome Consortium. Initially 18 Y-haplogroups were defined by YCC but later with the discovery of some more SNPs, their number was increased to 20, ranging from A-T. The frequency and occurrence of these haplogroups varies in different geographical regions and some haplogroups are found to be population specific as well.

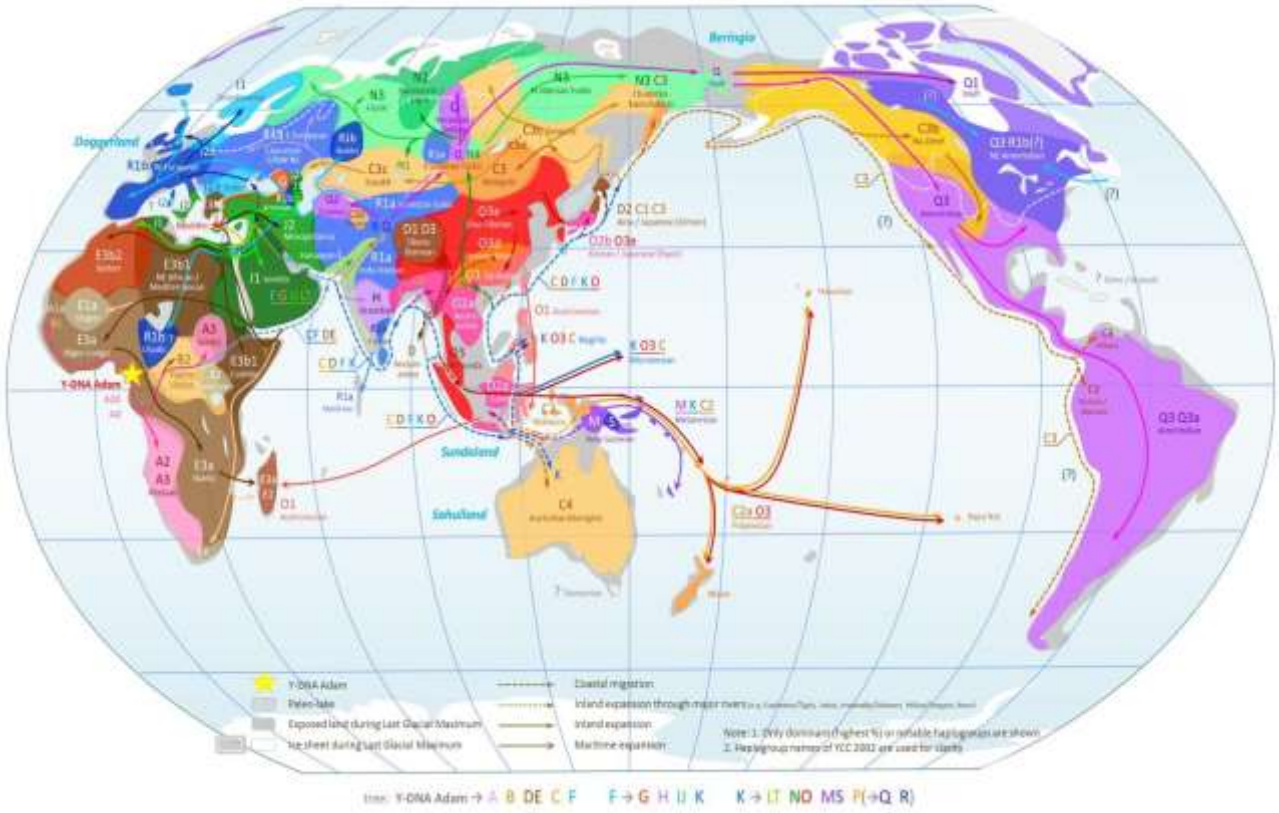



Fig.8. World map of Y- DNA Haplogroups around the globe with possible migration routes.(https://en.wikipedia.org)

### 2.7.4.1 Y-Haplogroups in South Asia

South Asia includes present day Pakistan, India, Afghanistan, Bangladesh, Nepal, Bhutan, Maldives and Sri Lanka. Several Y-DNA haplogroups including C, F, G, H, J, L, O, P, Q, R1a, R1b, R2 and T prevailed with varying frequencies in different ethnic groups inhabiting South Asia.

#### 2.7.4.1.1 Haplogroup R

Haplogroup R is said to be originated around 30,000 years ago in Central Asia as a branch of mega haplogroup P. There it divided into two branches, The R1a and R1b.

R1a moved towards Europe and R1b migrated towards South Asia. R1a is believed to be more likely instigated in South Asia and less probably in Eurasian steppes in the population of Kurgan culture practicing pit grave culture and horse domestication and the speakers of Indo European language. The high frequency of R1a haplogroup in south Asia is the consequence of Indo-Iranian migrations.

**2.7.4.1.2 Haplogroup Q**

The probable site of origin of this haplogroup is The South central Siberia or Altai Mountains around 17,000 to 31,700 years ago (Zegura *et al*., 2004; Sharma *et al*., 2007) and have been postulated to reach South Asia with Indo Aryan migrations. It is found in different populations of South Asian countries with varying frequencies.

**2.7.4.1.3 Haplogroup F**

This haplogroup has been found in south India and is considered to the parent of all other haplogroups (G-T) (fig.9). It prevails in almost 90% of the world" s population.Haplogroup G, H, IJ and K are considered to be the major sub-haplogroups of F. The probable place of its origin is either Eurasia around 48,000 years ago (Karafet *et al*., 2008) or South Asia. According to other study it has been originated in Levant or the Arabian Peninsula around 50,000 years ago (Hammer *et al*., 2002). Its descendant haplogroups depict pattern of expansion and radiation from South Asia or Middle East. And the presence of its sub clades in Africa might be the consequence of Back migration from south Asia or Southwest Asia towards Africa.

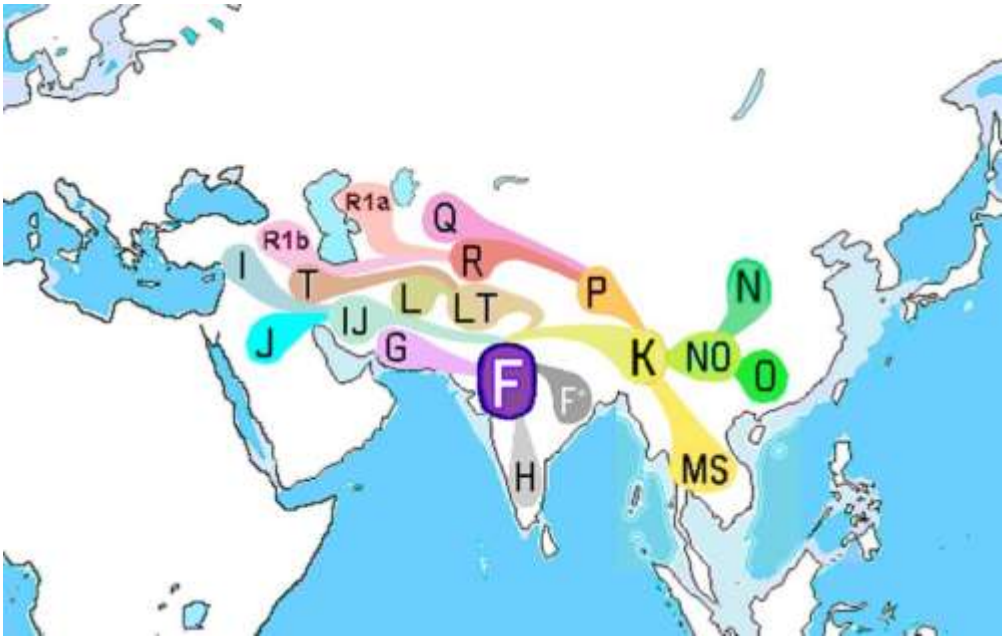



Fig.9. Y-DNA Haplogroup F and its descendent haplogroups from G to T (https://en.wikipedia.org)

**2.7.4.1.4 Haplogroup G, H, IJ**

Haplogroup G was found to be originated approximately17,000 - 21,000 years ago in the Middle East and is prevalent in several ethnic groups of Eurasia (Passarino *et al.*, 2002; Karlsson *et al.*, 2006) while haplogroup H is found primarily in South Asia. This haplogroup has been originated in South Asia around 30,000–40,000 years ago. Haplogroup I, is of European origin before Last Glacial Maximum from its ancestor haplogroup IJ which was said to be migrated there from middle East around 30,000-40,000 years ago. It bifurcated into I1 and I2 approximately 28,000 years ago (Battaglia *et al.*, 2009). Haplogroup J was derived from IJ in near east or West Asia around 42,000 years ago before present and it was scattered to India, Pakistan, Central Asia, Europe and

North Africa as a consequence of either Neolithic expansion or episodic migrations (Semino *et al*., 2004; Shou *et al*., 2010).

#### 2.7.4.1.5 Haplogroup L

This haplogroup first descended from haplogroup K approximately 30,000 years ago. It is found in high frequencies in various South Asian populations including Pakistani and Indian populations. Three sub branches of this haplogroup: L1, L2 and L3, have been found in Irani and Pakistani populations while only L1 has been found in Indian population (Sangupta *et al*., 2006).

#### 2.7.4.1.6 Haplogroup O

This haplogroup was originated in East Asia approximately 35,000 to 40,000 years ago and expanded towards South East Asian populations (Yan *et al*., 2011). Small percentage of this haplogroup is found in South Asian populations.

#### 2.7.4.1.7 Haplogroup T

Previously it was known as K2 (Mendez *et al*., 2011). It is said to be originated between Western and South Western Eurasian plate i.e. between Himalayas and Germania approximately 40,000 years ago (Hallast *et al*., 2015). According to several other investigations, K2 or T has been originated in Asia and later migrated to North Africa (Underhill *et al*., 2001, Cruciani *et al*., 2002; Samino *et al*., 2002; Luis *et al*., 2004). It is present in low frequency in South Asian populations.

The area under investigation includes the major area of Peshawar valley and is the best representation of Pathan population of Pakistan. The area is unexplored land with respect to human biology, particularly molecular anthropology. Hence the PhD research project presented here was started to establish a base line for molecular anthropology with the objectives highlights given in the introduction section.

## Chapter3

# MATERIALS AND METHODS

### 3.1 Materials and Methods

Buccal swab samples were obtained from 374 unrelated healthy male donors. 75 each from Muhammadzai, Mohmand, Kakakhel Mian populations from Charsada District, and Yousafzai and Mohmand (74 sample) from Mardan District. All the unrelated donors who voluntarily donated their samples after properly signing the informed consent were included in this study. For obtaining the oral swab volunteers were asked to rinse their mouth before giving sample, in order to minimize the chance of contamination. Each individual was provided a collection cup with having 3mL of 5% sucrose solution. The purpose of sucrose was to induce the production of Ptyalin carrying saliva, in the buccal cavity. Ultimately sucrose was digested and loose cheek epithelial cells were mixed with the saliva after rinsing their mouth with the solution for 2-2.30 minutes. The resulting solution was collected in the collection cup. Samples were then stored at -20 °C before further processing.

### 3.2 DNA Extraction

Genomic DNA was isolated from buccal swab by using phenol chloroform method as described by Marisi and Sergio (2007).

The solutions and reagents were prepared as protocol given in Appendix: II.

Samples were transferred to the labeled 1.5mL eppendorf tubes. These tubes were then centrifuged at 7000 rpm for 5 minutes to pellet the buccal cells and debris. The

supernatant was poured off immediately to avoid pellet slippage. This step was repeated twice or thrice to get maximum quantity of peletted cells.

In the next step, 300µL of cell lysis solution [10mM Tris, 0.5% SDS, 5mM EDTA (PH 8.0)] and 3 to 4 µL of Proteinase K (20 mg/ml) was added to each tube containing cell pellet. The mixture was vortexed at high speed for a few seconds to get the palette completely dissolved in the lysis buffer, and tubes were then incubated at 65°C for one hour. After incubation, 500 µL solution of Phenol chloroform in the ratio of 1:1 was added to each tube and was centrifuged at 10,000 rpm for 15 minutes. After that supernatant was transferred to the new eppendorf tube and 500 µL of Isopropanol (2propanol) was added and tubes were incubated for 10 minutes at -20°C. After that tubes were centrifuged at 12000rpm for 10 minutes and supernatant was discarded. 50 µL of

70% Ethanol was added to each tube and was centrifuged at 7000rpm for 5 minutes. This process was repeated after discarding the supernatant. Finally after palette washing with ethanol, tubes were kept inverted on clean absorbent paper overnight to let the palette dry. DNA palette was re-suspended in 50µl of TE buffer [12mM Tris (pH8) and 1mM EDTA] and was incubated for 10 minutes at 56°C to completely

dissolve the palette.

**3.3 Gel Electrophoresis**

The isolated DNA from all the samples was run on agarose gel for qualitative analysis.

1% agarose gel was prepared by adding 1g of agarose in 100mL of TAE- buffer. The mixture was then heated in the microwave oven until it boiled. Solution was then cooled to 40°C and 12 µL of Ethidium bromide was added to it. The solution was then poured

into the gel cassette with the combs suspended in it to create wells and was kept at room temperature until solidified. After solidification, gel was transferred to the gel tank containing running buffer (TAE- buffer). 5 µL of sample with 2 µL of loading dye was loaded to each well. It was then run for 20 minutes at 80 Volts. The gel bands were visualized under the UV light to confirm the presence of genomic DNA in the samples (fig.10).

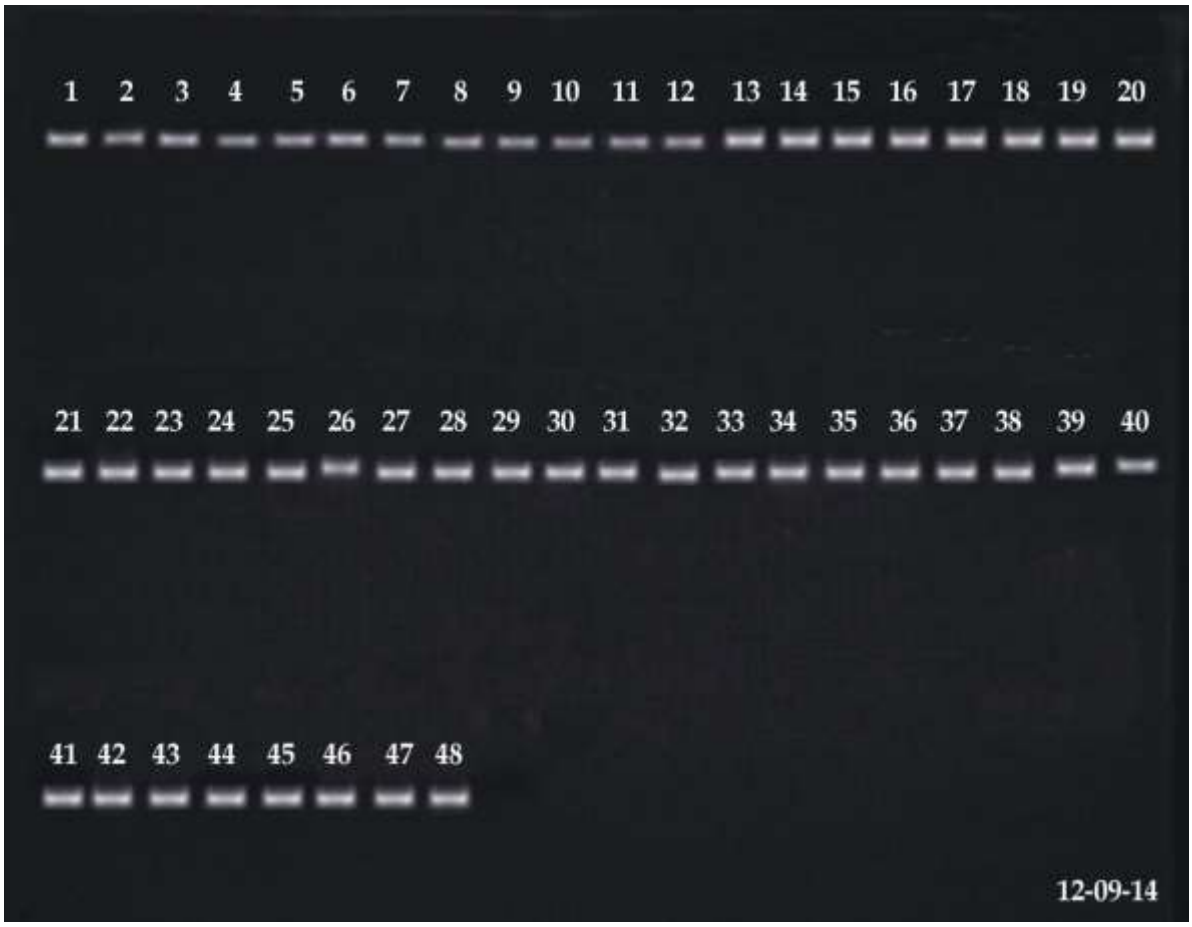



Fig.10. Agarose gel electrophoresis picture of the isolated Genomic DNA.

### 3.4 DNA Quantification

The DNA in each sample was quantified using Nano drop spectrophotometer before PCR amplification. And for those samples having highly concentrated DNA, serial dilutions were prepared before amplification.

### 3.5 Amplification of AMELY locus

All samples were amplified for AMELY locus confirmation before Y-STR amplification. 25µL of reaction mix was prepared for each 3 µL of template DNA (table.1) and all the samples were amplified using the thermo cycling conditions given in fig.11

Table:1. Components of AMELY PCR amplification.

| S. No | Components | Volume |
|---|---|---|
| 1 | Buffer | 2.5 µL |
| 2 | dNTPs | 1 µL |
| 3 | $MgCl_2$ | 1 µL |
| 4 | F primer | 1 µL |
| 5 | R primer | 1 µL |
| 6 | Plat Taq | 0.2 µL |
| 7 | 5% Glycerol | 5 µL |
| 8 | Water, Amplification Grade | 13.3 µL |
| | Total volume of master mix | 25 µL |

#### 3.5.1 Thermo cycling parameters of the PCR

Thermal cycler was programmed with the following conditions as summarized in Fig.11.

i. Initial hot start at 94°C for 2 minutes ii.

Denaturation 94°C for minute (35

cycles) iii. Annealing at 60°C for 1 minute

iv. Extension at 72°C for 5 minutes

v. Final extension was performed at 72°C for 5 minutes

vi. Storage soak indefinitely at 4°C

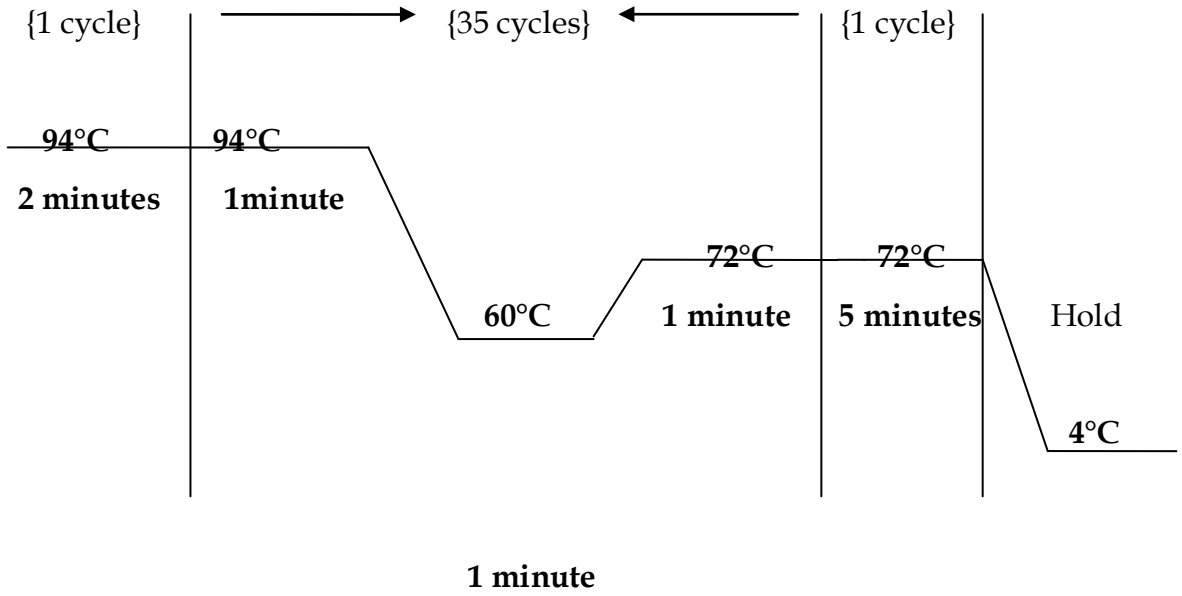



Fig.11. Thermo cycling conditions of AMELY amplification.

### 3.6 Polyacryleamide gel electrophoresis (PAGE)

All the PCR products were run on Polyacryleamide gel to confirm the efficiency of PCR amplification. Polyacryleamide gel was prepared by adding 4.8mL $dH_2O$, 0.65mL 10X TBE, 975 µL of 40% Acryl mix, 16µL of 25% APS and 16µL of TEMED.

All the samples except two were found AMELY positive for further STR amplification with no risk of deletion of EMELY locus along with four important Y-STR loci i.e.

DYS570, DYS456, DYS576 and DYS481. A representation picture of PAGE results is provided in Fig.12.

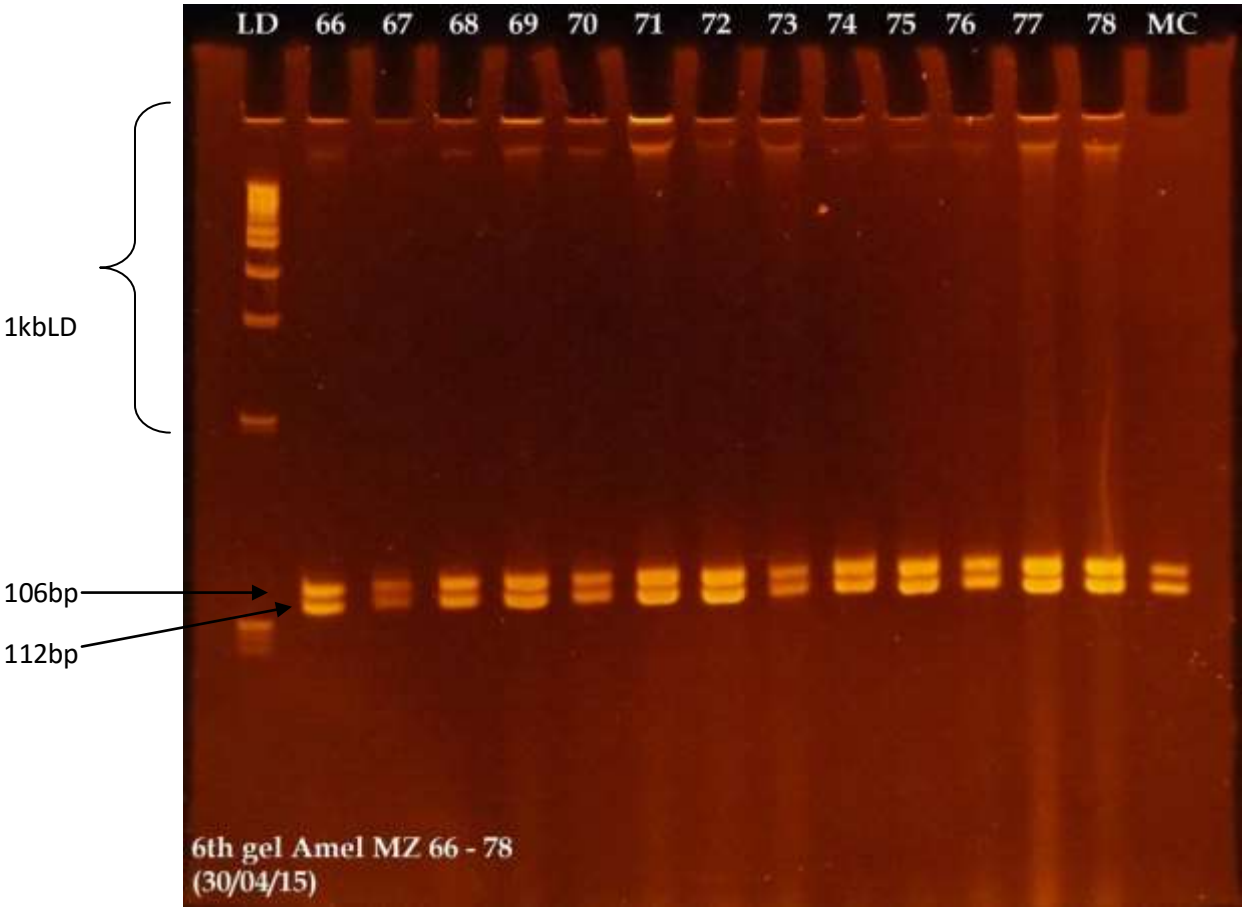



Fig.12. Polyacryleamide Gel Electrophoresis photograph of AMELY PCR product. LD: Ladder and MC: Male control

**3.7 Y-STR amplification**

Total of 374 samples from five different Pathan populations of Charsada and Mardan district were amplified for 23 Y chromosomal loci including (DYS576, DYS389I, DYS448, DYS389II, DYS19, DYS391, DYS481, DYS549, DYS533, DYS438, DYS437, DYS570, DYS635, DYS390, DYS439, DYS392, DYS643, DYS393, DYS458, DYS385ab, DYS456,

GATAH4 ) using Promega PowerPlex® Y23 System (PPY23, Promega Corporation, Madison, WI). Relative positions of 23 STRs on Y chromosome are provided in Fig.13.

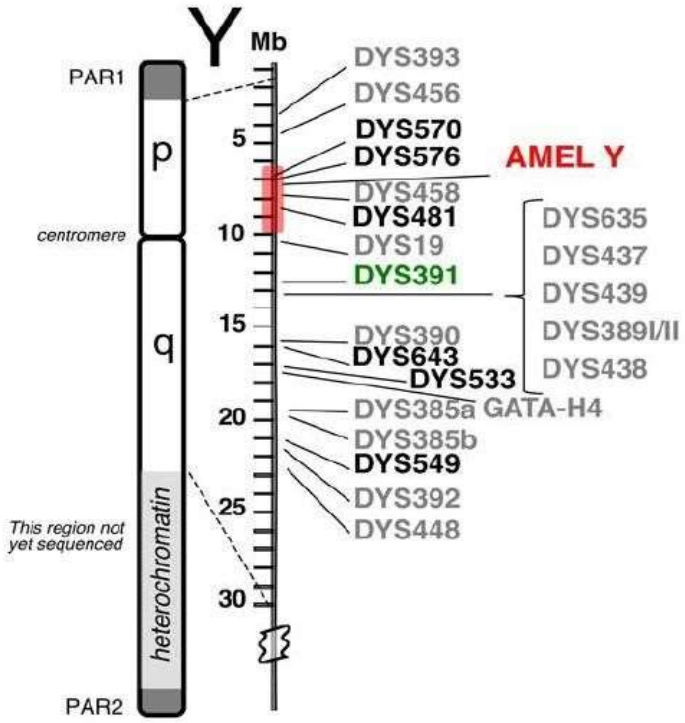



Fig.13. Relative positions of the 23 Y-STR loci used in the Promega PowerPlex® Y23 System.

### 3.7.1 Procedure

All the Pre-amplification components were thawed just prior to use. All the tubes were centrifuged and vortexed for 15 seconds before each use. Number of reactions was determined, including positive and negative control and 1 or 2 reactions were added to that number. PCR amplification mix was prepared by combining all the components of the kit using protocol established by Promega PowerPlex® Y23 System given in the table.2.

PCR Amplification mix was vortexed for 5-10 seconds and after that 9.4 µL of PCR was transferred to each labeled 0.2mL tubes. To each tube containing PCR amplification Mix, 1.5 µL of 0.5ng/µL DNA template was added.

Table: 2. Components of Promega PowerPlex® Y23 amplification system.

| **Component** | **Volume per Reaction** |
|---|---|
| Promega PowerPlex® Y23 5X. Master Mix | 1.25 µL |
| Promega PowerPlex® Y23 10X Primer Pair Mix | 0.625 µL |
| $MgCl_2$ | 0.6 µL |
| Water, Amplification Grade | 7 µL |
| Total Volume of Amplification Mix. | 9.475 µL |
| Template DNA | 0.5ng/ µL of 0.5 µL |
| Total Reaction volume | 11 µL |

### 3.7.2 Thermo cycling parameters of the PCR

Thermal cycler was programmed with the following conditions (Fig.14).

i. Initial hot start at 96°C for 2 minutes
ii. Denaturation 94°C for 10 seconds (28 cycles)
iii. Annealing at 61°C for 1 minute iv. Extension at 72°C for 30 seconds
v. Final extension was performed at 60°C for 20 minutes
vi. Storage soak indefinitely at 4°C

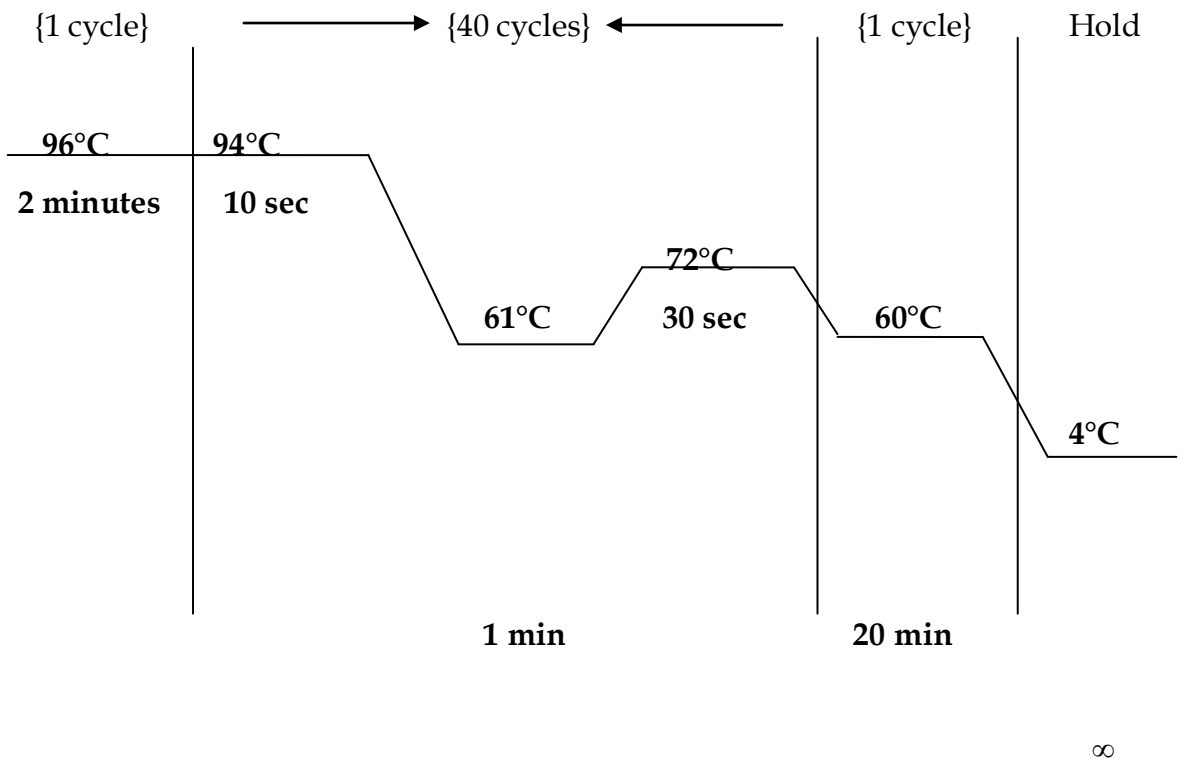



Fig.14. Thermal cycling protocol for PPY[23] System using Eppendorf mastercycler ep Gradient S thermal cycler.

### 3.8. Polyacryleamide Gel Electrophoresis

The next step was to check to make sure the PCR worked, before moving on to the next step. For this, Polyacryleamide gel electrophoresis (PAGE) was used. (Fig.15)

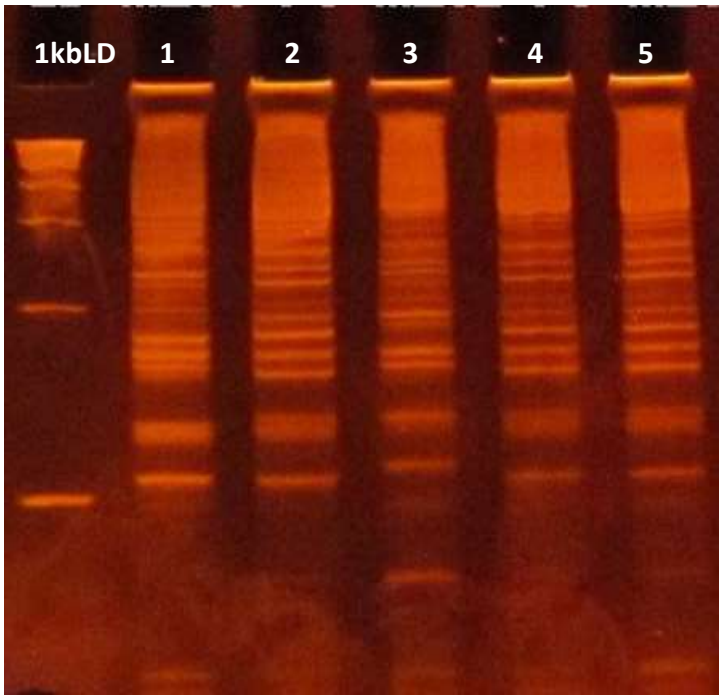



Fig.15. Polyacryleamide gel electrophoresis photograph of 23 Y-STR loci amplified with PPY23 System.

### 3.9 Capillary Electrophoresis

After confirmation, capillary electrophoresis was carried out for fragment analysis using Applied Biosystem 3730 Genetic analyzer.

#### 3.9.1 Sample Preparation

Sample preparation for ABI 3730 Genetic Analyzer was carried out initially by diluting each of the 5 dye matrix standard (Flourescein, JOE, TMR-ET, CXR-ET and CC5) in the ratio of 1:10 with Nuclease free water after thawing them. After this Fragment mix was prepared.

15 µL of each of the 5 dye matrix standards was taken from the initial dilution and was mixed with 1425 µL of Hi-Di™ formamide. 96 wells were used on 96 capillaries for matrix detection on ABI PRIMSM® 3730 Genetic Analyzer. 25 µL of fragment mix was loaded to each well containing 1 µL of sample in each well and 1 µL of PowerPlex® Y23 allelic

ladder mix to one of the wells, and plate was centrifuged for a few seconds to remove bubbles.

PCR product was denatured at 95°C for 3 minutes, and then was immediately chilled on crushed ice for 3 minutes, just before loading to the instrument.

### 3.9.2 Reading Alleles calls

The peak sizes and positions of the amplified alleles were compared with the PowerPlex® Y23 allelic ladder using Gene Mapper® Software 5 after capillary electrophoresis. The range of different alleles for 23 Y-STR loci is given in the fig.22.

### 3.10 Comparative populations

To investigate the genetic diversity of five major Pathan tribes of Charsada and Mardan districts and to compare with the other Pathan populations in the neighboring area of Mohmand agency and Swat valley and populations from neighboring countries, their Y chromosome genotype data was exploited.

For Y STR analysis, the comparative population data used was obtained from the published literature of Afghanistan[Pathan] (Lacau *et al*., 2012), Iraq[Iraqi] (Purps *et al*., 2014), Iran [Arabs, Bakhtiari, Galaki, Mazandarani, SouthernTalysh] (Rower *et al*. 2009), Turkey [Dogukoy, Eskikoy, Gocmenkoy, Merkez] (Alakoc *et al*., 2010), Greece [Greeks] (Purps *et al*., 2014), Israel [Muslim Arabs] (Farnandes *et al*., 2011), Russian Federation [Archangelskaja, Brianskaja, Ivanoskaja, Lipezkaka, Ryasankaja, Smolenjkaja, Tamboskaja, Tverskaja, Wologodskaja] (Rower *et al*., 2008), India [Tamil] (Balamurugan *et al*., 2010), Xinjiang China [Uighur, Kazakh] (Shan *et al*., 2014), Tibet China [Tibetan]

(Ye *et al*., 2015), Pakistani Pathans (Lee *et al*., 2014) and Yousafzai Pathans (KP) Pakistan (Ilyas *et al*., 2012).

### 3.11 Mitochondrial DNA amplification

Out of 374 samples 165 samples from all five populations were amplified with 15971F (TTAACTCCACCATTAGCACC) and 484R (TGAGATTAGTAGTATGGGAG) primers and the sequence length of 1082bps was amplified using 25 µL total reaction volume. The amplification was carried out using" Eppendorf Mastercycler ep Gradient S thermal cycler" of Applied Biosystem (Fig.16).

Recipe is given in the table: 3.

Table: 3. Reagents used in PCR reaction mixture

| S. NO | Reagents | Volume of Reagents |
|---|---|---|
| 1 | 10X buffer w/o $MgCl_2$ | 2.5 µL |
| 2 | 10mM dNTPs | 1 µL |
| 3 | 50mM $MgCl_2$ | 1 µL |
| 4 | (5 µM) F Primer | 1 µL |
| 5 | (5 µM) R Primer | 1 µL |
| 6 | 5% Glycerol | 1 µL |
| 7 | Platinum Taq polymerase (5U/ µL) | 0.2 µL |
| 8 | $ddH_2O$ | 14.3 µL |
| 9 | Template DNA | 3 µL |
| | Final Volume | 25 µL |

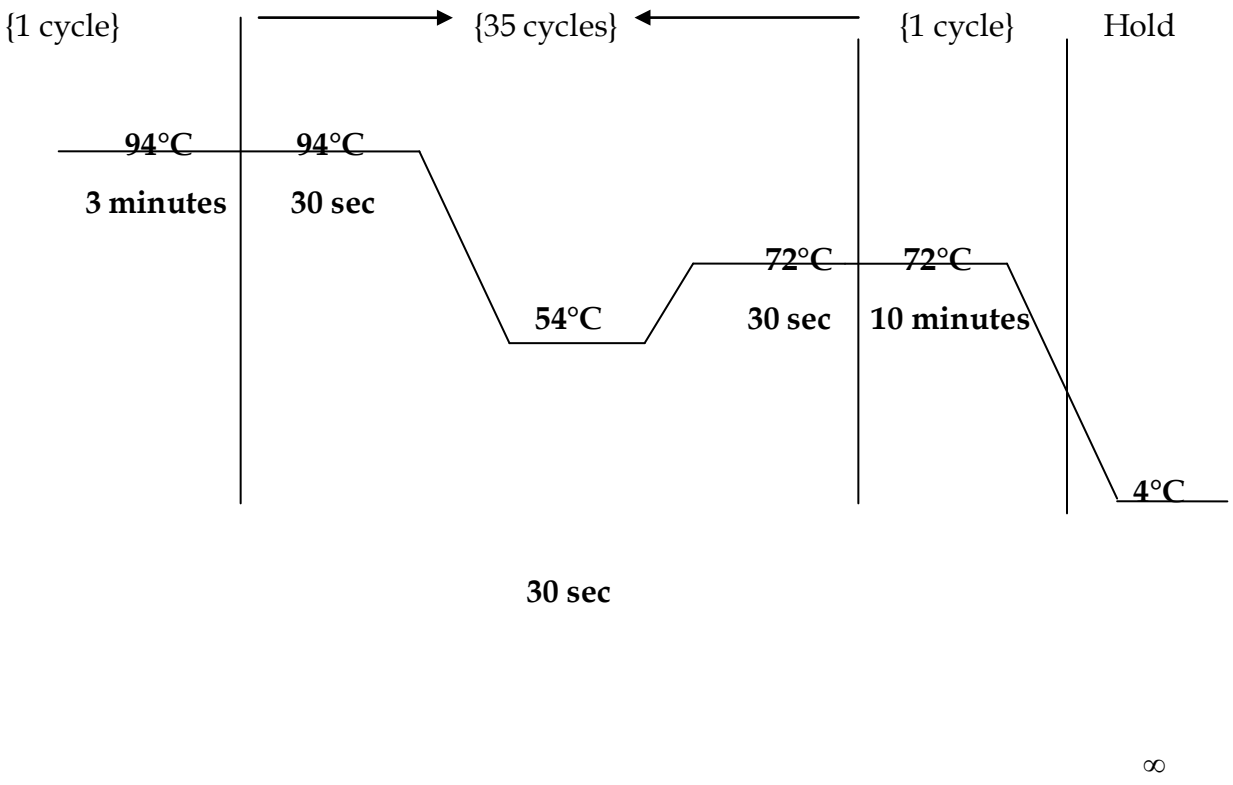



Fig.16.Thermo cycling conditions for mtDNA amplification

**3.12 PCR product clean-up**

5µL of the PCR product was mixed with 2 µL of Exo-SAP-it in minitubes (SAP: Shrimp Alkaline Phosphatase and EXO: Exonuclease I) (Malhi *et al*., 2010). After spinning, tubes were incubated at 37°C for 30 minutes followed by heat inactivation at 80°C for 15 minutes and were held at 25°C (Fig.17).

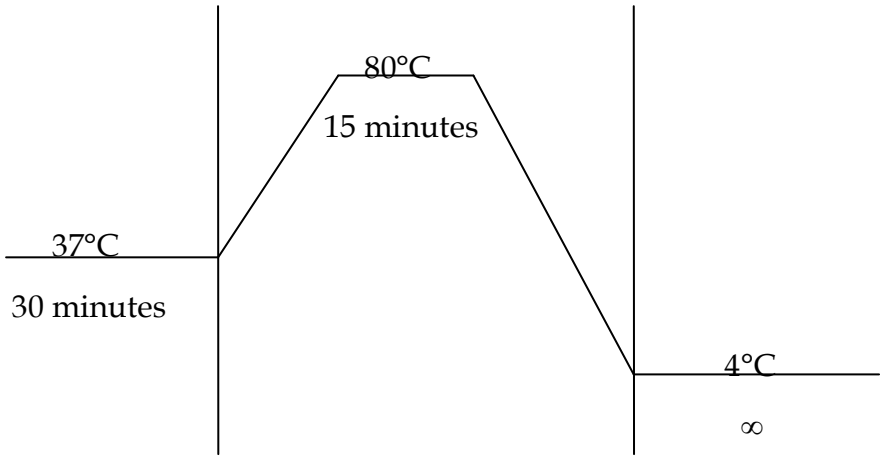


Fig.17. Thermo cycling conditions for EXO- SAP.

The purpose of this procedure was to chew up excess primers and removal of excess dNTPs from the PCR product, to make the DNA sequences readable.

**3.13. Agarose Gel Electrophoresis**

In order to visualize the PCR product to make sure that it worked and there were no multiple bands, the PCR product was electrophorased in 2% agarose gel. In order to prepare 2% agarose gel, 3g of agarose was added to 150mL of 1X TBE buffer. Mixture was boiled in microwave oven for 1.30 minutes at the interval of 30sec in order to avoid spillage. Solution was then cooled up to 45°C. After sufficient cooling Ethidium bromide was added in to it and solution was poured in to the gel cassette with the combs fitted into it to create wells. After solidification, combs were removed and 5 µL of each PCR product was mixed with 2 µL of loading dye. All the samples were loaded in the separate well. 2-Log DNA ladder (0.1 – 10.0kb) was loaded in the first well for comparison. Representation of Gel results is given in Fig.18.

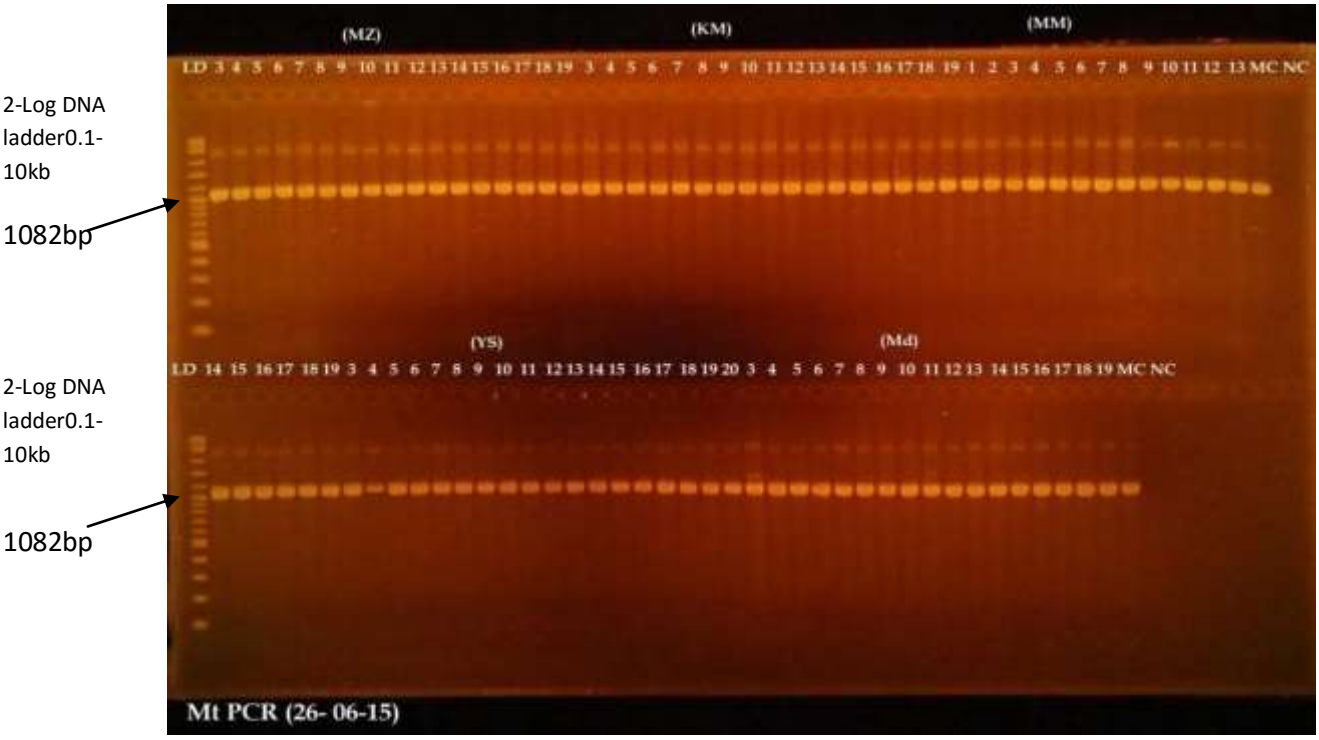



Fig.18. Agarose Gel Electrophoresis photograph of amplified mtDNA. MC: Male control, NC: Negative control

### 3.14 mtDNA Sequencing

For mitochondrial DNA Hyper variable region I (HVI) sequencing, the 2 µL of good quality clean PCR product was mixed with 1 µL of 5 µM of 15971 F primer.Sequencing was carried out using Big Dye™ terminator cycle sequencing kit of Applied Biosystem on Genetic analyzer 3730.

#### 3.14.1 Sequence reading

After retrieval, Sequences were read using the program Sequencher 4.7 and all sequence data was aligned with revised Cambridge reference sequence (rCRS) (rCRS; Andrews *et al*., 1999) using bioedit program. Variable positions were determined from np 16024 – 16365 and haplogrup was assigned to each sample using phylotree build 16 (Van Oven M.2009) and mitomaster (Lott *et al*., 2013).

To investigate the genetic diversity of five major Pathan tribes of Charsada and Mardan districts and to compare with the other Pathan populations in the neighboring area of Mohmand agency and Swat valley and populations from neighboring countries, their mtDNA HVI sequence data was exploited.

### 3.15 Comparative populations

For mtDNA analysis the comparative data from populations of India [Afridi, Tamil] (Metspalu *et al*., 2004), Israel [Druze, Bedauin] (Hunley *et al*., 2009), Greece [Greeks] (Irwin *et al*., 2008), Russian Federation [Yakut] (Hunley *et al*., 2009), Iran [Armenian, Azeris, Persian, Qashqis] (Derenko *et al*., 2013), Iraq [Iraqi] (Abu *et al*.,2008), China [Uighur] (Hunley *et al*., 2009), Afghanistan, Turkey [Turkish] (Comas *et al*., 1996), indigenous populations from Pakistan [Pathan, Kalash, Brusho, Baluch, Brahui] (Hunley *et al*., 2009) was used.

### 3.16 Statistical analysis

Genetic diversity estimates were calculated for all the five Pathan populations of the two areas separately as well as collectively using a formula [$n(1-\sum p_i^2)/n-1$], where $n$ is the number of samples and $p_i$ is the frequency of the $i^{th}$ haplotype in the population (Nei. 1987). Genetic distances (Matrix of Slatkin linearized $F_{ST}$) were obtained by using ARLEQUIN Ver 3.5.2.2 software for our data and for other global populations. Of all pairwise $F_{ST}$ values a distance matrix was constructed that was later used to construct a PCoA plot using GENALEX program (Peakall *et al*., 2006). Median Joining (MJ) networks were constructed using NETWORK 4.6.1.3 (http://www.fluxus-

engineering.com) by giving each locus a weight according to its estimated mutation rate (Keyser *et al*., 2003). Mitomaster and phylotree were used to predict the mitochondrial haplogroups (http://www.mitomap.org) (Lott *et al*., 2013) and Y-haplogroups were identified by using server haplogroup predictor (www.hprg.com).For STR analysis, the haplotypes were reduced to 17 STR loci (DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, Y-GATA-H4) for comparison with published available global populations data sets.

**3.17 Mantel test**

At the end Mantel test were performed using GENALEX program (Peakall *et al*., 2006) to measure the degree of correlation between mtDNA based Genetic distance among population and geographical distance, between Y-STR based Genetic distance and Geographic distance and between mtDNA based and Y-STR based genetic distances.

## Chapter4
# RESULTS

This chapter includes Y-STR and mtDNA analysis results profiles consisting of Genetic Diversity parameters for Y-STR and mtDNA HVI data sets, Haplogroup analysis, Network analysis, Analysis of Molecular Variance, Principal Coordinate analysis and Mantel Test.

### 4.1 The Y-STRs analysis

Analyses of the results obtained from Y-STR DNA samples obtained from 374 volunteers of the districts Mardan and Charsada segregated into 344 haplotypes wherein 374 samples were belonging only to 11 haplogroups (Annexure I), with an overall Discrimination Capacity (DC) of 92%.

### 4.2 Genetic Diversity

The genetic diversity (GD) obtained for all the five Pathan populations was 0.99 which is similar to the already studied Pathans from Pakistan (Lee *et al.,* 2014). Genetic diversity estimatesand other population genetics parameters for different tribes are given in the table 4.Comparison of genetic diversity estimates among various indigenous Pakistani populations is given in table 5.

Table: 4. Population genetics parameters for 23 Y-STRs in five Pathan populations of Charsada and Mardan district.

| Parameters | CHARSADA | | | MARDAN | | All 5 populations |
|---|---|---|---|---|---|---|
| | (MZ) | (MM) | (KM) | (YS) | (MD) | |
| **Number of samples(n)** | 75 | 75 | 75 | 75 | 74 | 374 |
| **No. of Haplotypes** | 72 | 71 | 61 | 73 | 67 | 344 |
| **Unique Haplotypes** | 70 | 67 | 51 | 71 | 62 | 320 |
| **Shared Haplotype** | 2 | 4 | 10 | 2 | 5 | 24 |
| **Random Match Probability (RMP)** | 0.0136 | 0.014 | 0.019 | 0.0133 | 0.0158 | 0.00305 |
| **Power of Discrimination (PD)** | 0.986 | 0.985 | 0.980 | 0.986 | 0.984 | 0.996 |
| **Genetic Diversity(GD)** | 0.999 | 0.998 | 0.994 | 0.999 | 0.997 | 0.998 |

Note: The abbreviations MZ, KM, YS, MM and MD denote the tribes Muhammadzai, Kakakhel Mian, Yousafzai and Mohmand respectively.

Table: 5. Comparison of Y-STR Genetic diversity among various ethnic groups of Pakistan

| **Parameters** | **Path (present study)** | **Mkr (Qamar *et al*. 2002)** | **Pth (Lee *et al*. 2014)** | **Bal (Qamar *et al*. 2002)** | **Brh (Qamar *et al*. 2002)** | **Haz (Qamar *et al*. 2002)** | **Bsh (Qamar *et al*. 2002)** | **Ksh (Qamar *et al*. 2002)** | **Par (Qamar *et al*. 2002)** | **Snd (Qamar *et al*. 2002)** |
|---|---|---|---|---|---|---|---|---|---|---|
| **No. of Sample** | 374 | 33 | 230 | 59 | 110 | 23 | 94 | 44 | 90 | 122 |
| **No. of haplotypes** | 344 | 30 | 211 | 48 | 85 | 11 | 63 | 26 | 60 | 101 |
| **Genetic Diversity** | 0.998 | 0.992 | 0.997 | 0.988 | 0.973 | 0.893 | 0.987 | 0.952 | 0.974 | 0.995 |

Note: The abbreviations Path, Mkr, Pth, Bal, Brh, Haz, Bsh, Ksh, Par, Snd are Pathan from Charsada and Mardan, Makrani, Pathan, Baluch, Brahui, Hazara, Barusho, Kalash, Parsi and Sindhi, respectively

### 4.3 Y-Haplogroups found

Analysis of haplogroups segregated from the DNA samples revealed that around 49.2% of the individuals belonged to one of three major Haplogroups i.e R1a, 17.9% from L haplogroup and 9.6% from G2a haplogroup. A relative incident of the haplogroups is presented in the table. 6.

Table: 6. Frequency of Y-haplogroups found in five populations under study and Pathan population from Southern Afghanistan.

| | **MZ n=75** (This study) | **KM n=75** (This study) | **MM n=75** (This study) | **YS n=75** (This study) | **MD n=74** (This study) | **Total N=374** (This study) | **Southern Afghanian (N=145)** (Sangupta *et al*. 2007) |
|---|---|---|---|---|---|---|---|
| **Haplogroup** | **Frequency of Haplogroups (%)** | | | | | | |
| **R1a** | 50.7 | 41.3 | 54.6 | 54.6 | 44.6 | 49.2 | 66.8 |
| **L** | 17.3 | 18.7 | 10.7 | 17.3 | 25.7 | 17.9 | 5.6 |
| **G2a** | 10.7 | 4 | 14.7 | 10.7 | 8.1 | 9.6 | 8.2 |
| **Q** | 8 | 13.3 | 4 | 5.3 | 5.4 | 7.21 | 1.3 |
| **H** | 8 | 10.7 | 1.3 | - | 1.4 | 4.27 | 2.8 |
| **J1** | - | 6.7 | - | 1.3 | 8.1 | 3.20 | 2.0 |
| **R1b** | - | 1.3 | 6.7 | 4 | 4 | 3.20 | 0.6 |
| **J2b** | 1.3 | 2.7 | 4 | - | 1.3 | 1.87 | - |
| **E1b1a** | - | 1.3 | - | 5.3 | - | 1.33 | 0.6 |
| **E1b1b** | 2.7 | - | 1.3 | - | 1.4 | 1.06 | 0.6 |
| **I2a** | 1.3 | - | 2.7 | 1.3 | - | 1.06 | - |

*(Note: The abbreviations MZ, KM, MM, YS, MD denotes the tribes Muhammadzai, Kakakhel Mian, Mohmand from Charsada, Yousafzai and Mohmand from Mardan, respectively).*

Haplogroups Q, H, J2b, J1, J2a, R1b, E1b1a and E1b1b were also present comprising 7.2%, 4.2%, 1.87%, 3.2%, 1.33% and 1.06% respectively of the entire sample set investigated. Of the haplotypes generated, 7% of all were shared among populations and almost 93% of the haplotypes were unique to the individuals.

Muhammadzai (MZ), Kakakhel Mian (KM) and Mohmand (MM) samples came from Charsada area and Mohmand (MD) and Yousafzai (YS) from Mardan area. Total of 7% of the haplotypes were shared among all the five populations of the two areas.

### 4.4 Network analysis of Y Haplogroups

The network analysis of the haplogroups showed that the Haplotype „a" is the central haplotype with overall frequency of 26.1% present in 3.8% of MZ haplogroup R1a haplotypes, 5.4% of MM Hg R1a haplotypes, 7.1% in KM R1a haplotypes, 6% in MD R1a haplotypes and 3.8% in YS R1a haplotypes. Similarly haplotype b is the second haplotype shared among all the five populations with the frequency of 6.5% in MM, 1.6% in MZ and MD and 0.5% in KM R1a haplotypes. Little haplotype sharing was observed in haplogroup H and G2a (Fig.19 a).

No common or shared haplotype was observed among the samples of either of the population in Haplogroup H while single haplotype was shared between the individuals of Mohmand from Charsada (MM), Kakakhel Mian (KM) and

Muhammadzai (MZ) in Haplogroup G2a (Fig.19 b, c).

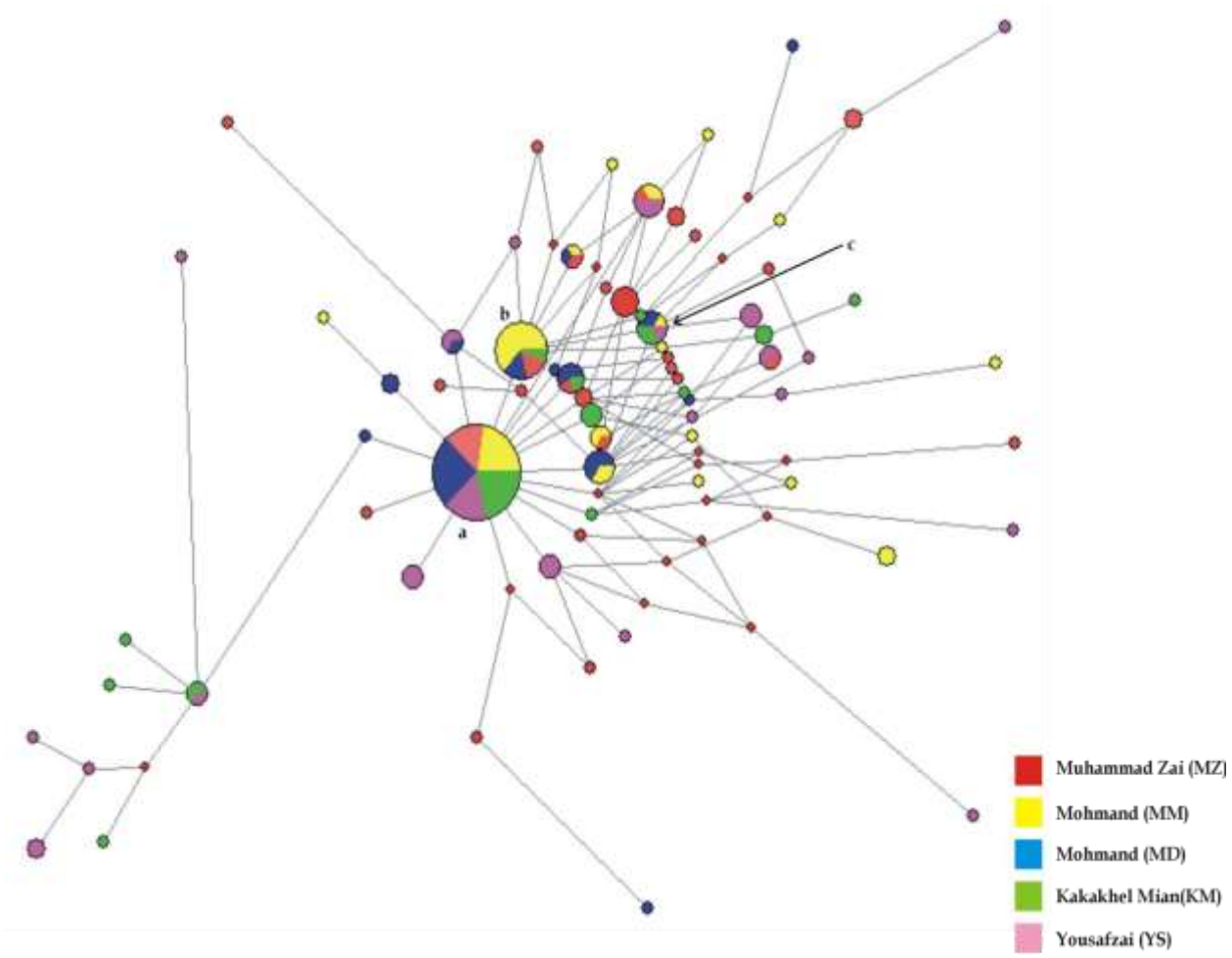



Fig.19.a. Median-Joining network of Y-haplogroup R1a in five populations of Charsada and Mardan. Areas of circles are proportional to the haplotype frequencies. Haplotype "a" and "b" are most frequent haplotypes shared among all the five populations.

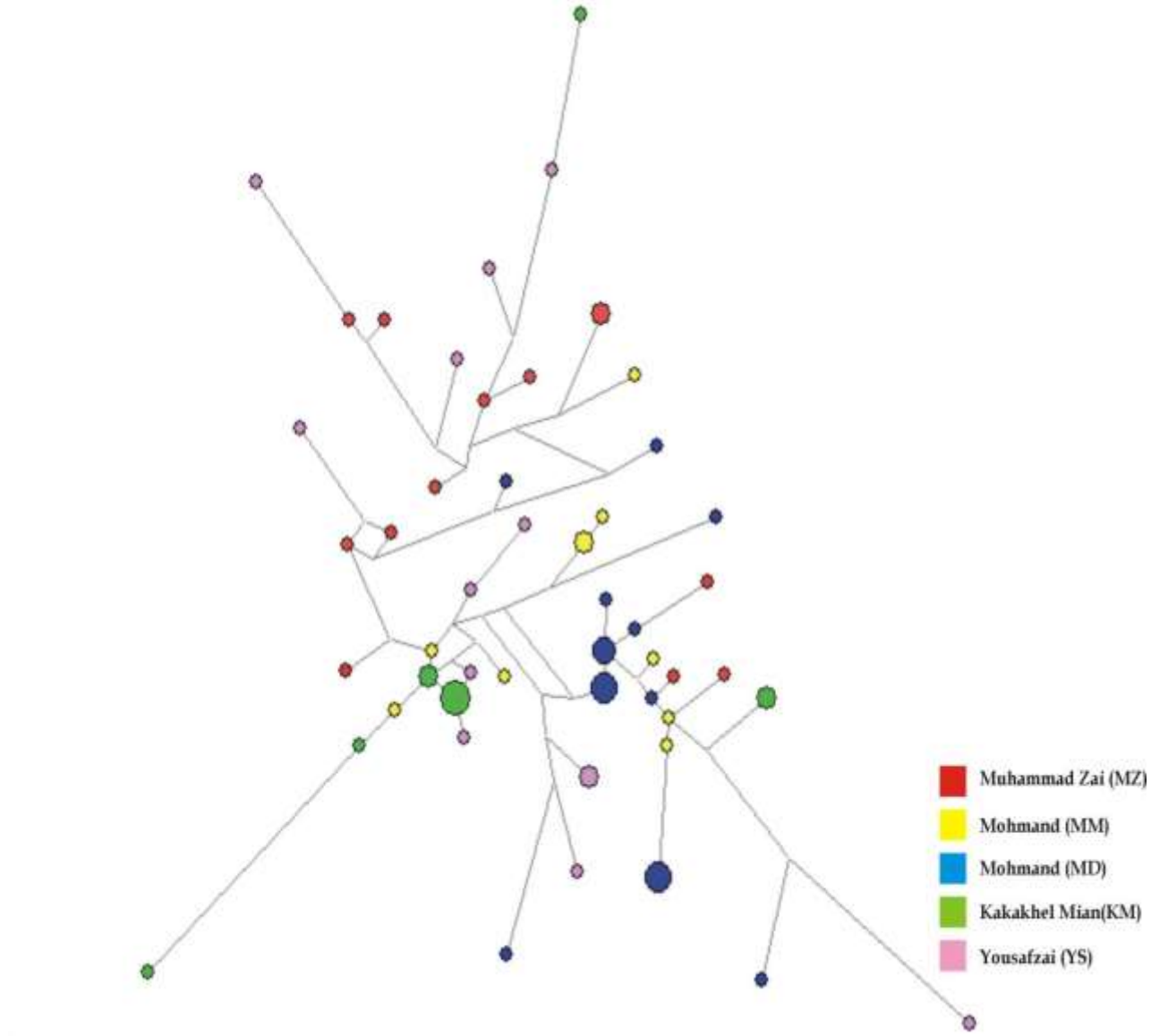



Fig.19.b. Median-Joining network of Y-haplogroup L in five populations of Charsada and Mardan. Areas of circles are proportional to the haplotype frequencies.

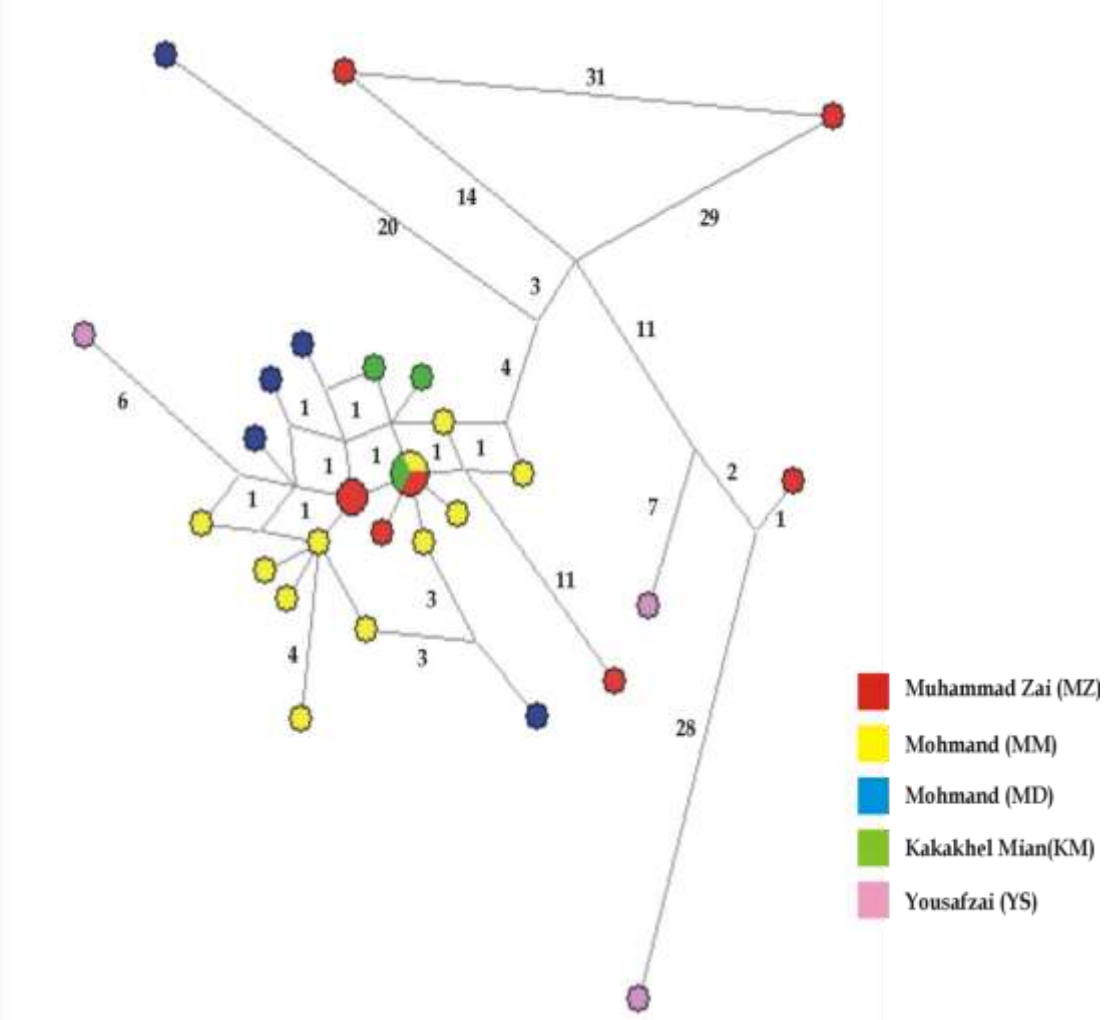



Fig.19.c. Median-Joining network of Y-haplogroup G2a in five populations of Charsada and Mardan. Areas of circles are proportional to the haplotype frequencies.

### 4.5 Haplogroup distribution among the population of two districts

Charsada populations had a few additional haplogroups not found in Mardan populations. More diverse haplogroups points out the different history of this location (Fig.20).

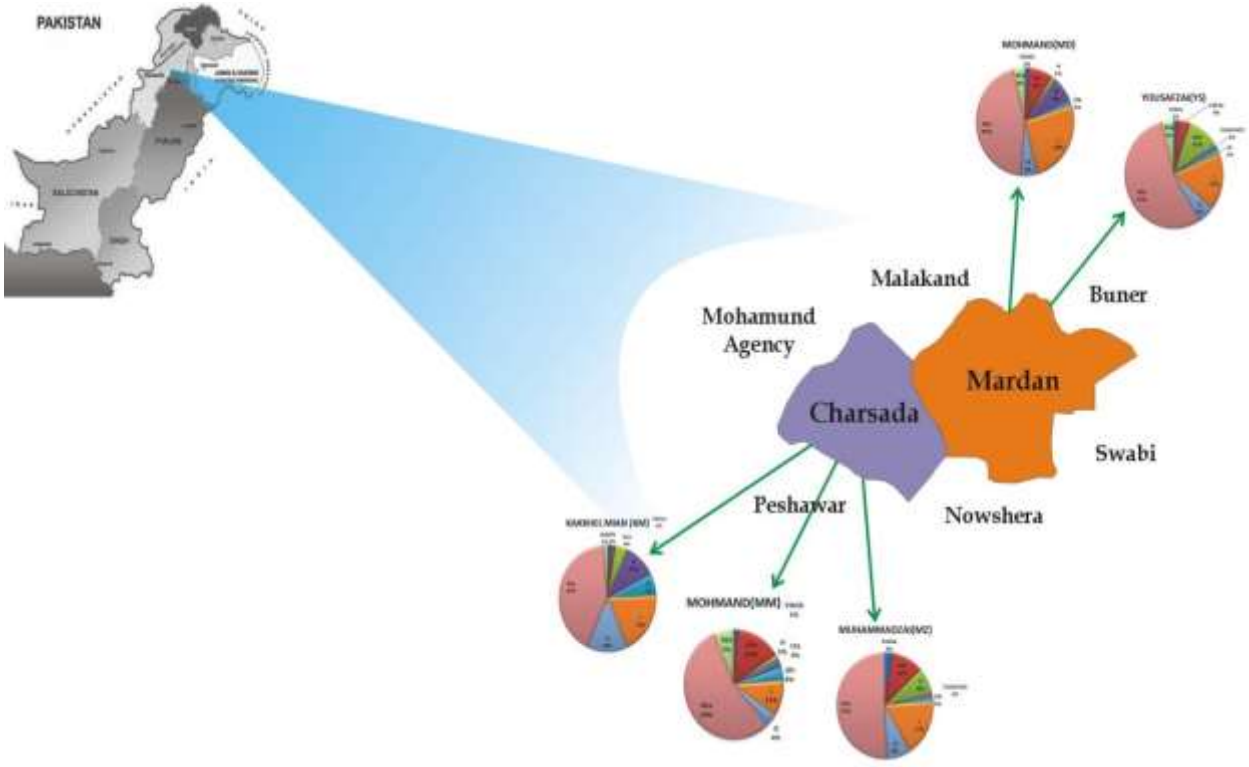



Fig.20. Y-Haplogroup distribution in five Pathan populations of Charsada and Mardan district.

### 4.6 Genetic distance among five populations

No significant genetic distance ($F_{ST}$) was found between the pair of populations from both the districts. The highest genetic distance of 0.04 between Kakakhel Mian (KM) and Mohmand (MD) was observed followed by 0.03 between Kakakhel Mian(KM) and Muhammadzai (MZ), 0.02 between Mohmand (MD) from Mardan and Muhammadzai (MZ) while no distance was observed between Muhammadzai, Mohmand (MM) from Charsada and Yousafzai (YS) Fig.21.

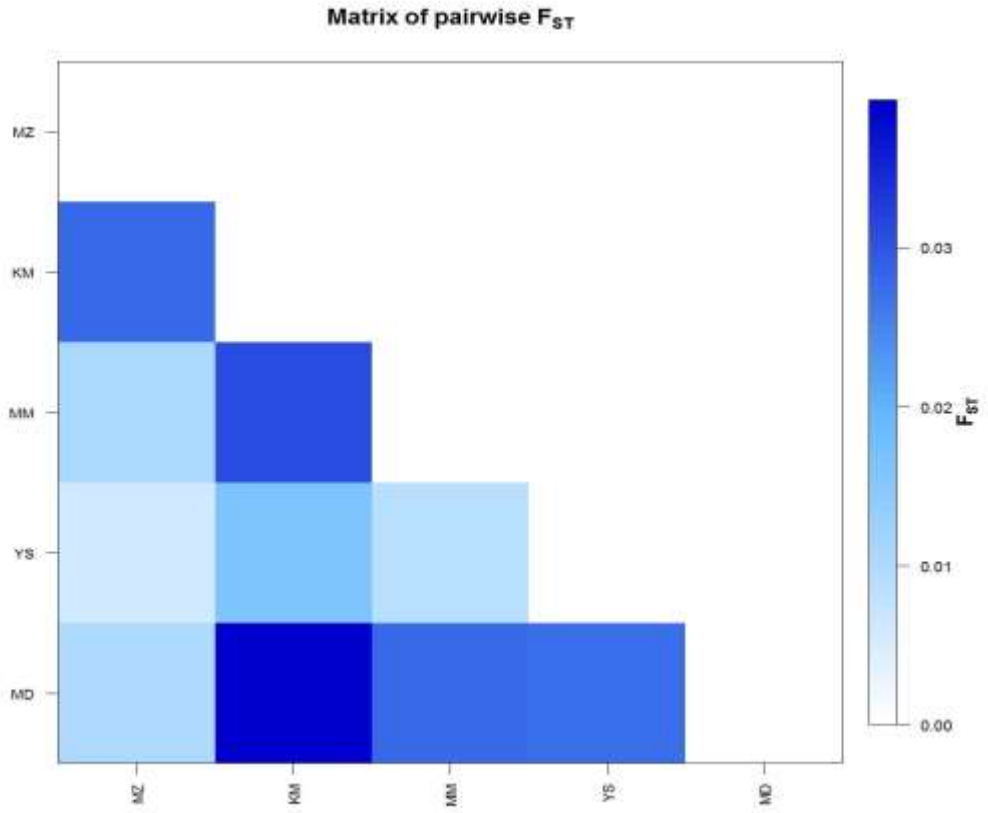



Fig.21. Y-STR based Pair wise Genetic distances ($F_{ST}$) between five Pathan population pairs of Charsada and Mardan district. *(The abbreviations MZ, KM, MM, YS, MD denote the tribes Muhammadzai, Kakakhel Mian, Mohmand from Charsada, Yousafzai and Mohmand from Mardan, respectively).*

### 4.7Analysis of Molecular Variance among population of two districts

AMOVA analysis for Y-STR haplotypes was used to examine the partitioning of genetic variation for these five populations of two districts divided into two groups. In this case 98.20% of variation found "within the populations" while -0.6% of variation was found "among the groups" and 2.46% "among population within the group" (table: 7).

Negative values indicate zero value or no variation between the groups.

Table: 7. AMOVA design and results (average over 23 Y-STR loci) in five populations of two districts (Groups).

| Source of Variation | Sum of Squares | Variance Components | Percentage Variation |
|---|---|---|---|
| **Among groups** | 13.379 | -0.04132 | -0.56909 |
| **Among populations within groups** | 62.417 | 0.18488 | 2.54624 |
| **Within populations** | 2601.665 | 7.11737 | 98.02285 |
| **Total** | 2677.460 | 7.26093 | |

## 4.8 Allele Frequencies in all the five Populations

Number of alleles at 23 different loci in all the five populations is graphically represented in fig.22.

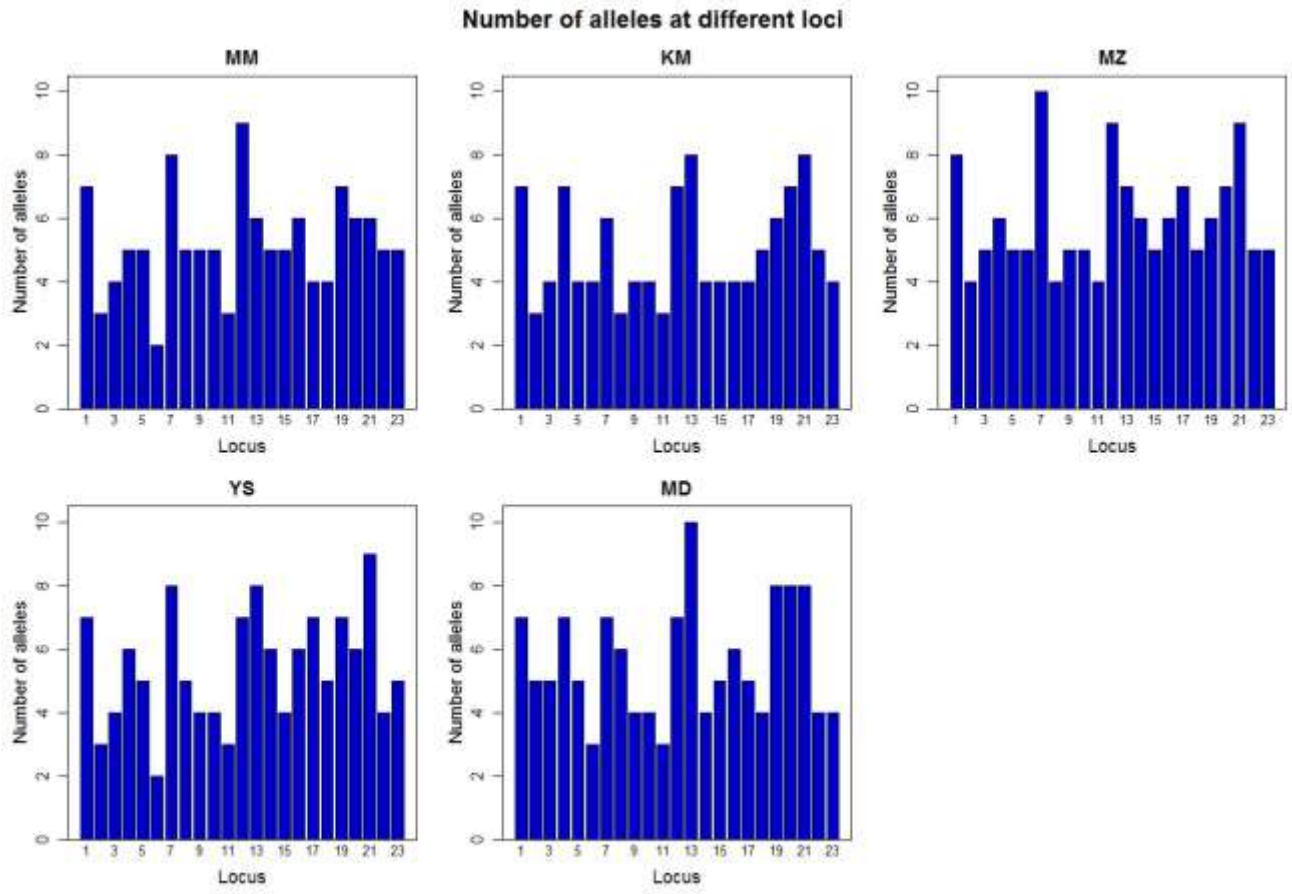



Fig.22. Graphical representation of allele frequencies of 23 loci in five populations of Charsada and Mardan district.
*(The abbreviations MZ, KM, MM, YS and MD denote the tribes Muhammadzai, Kakakhel Mian, Mohmand from Charsada, Yousafzai and Mohmand from Mardan, respectively).*

**4.8**

**.1 Yousafzai: YS (Mardan)**

Allele frequencies of 23 Y-STR loci in Yousafzai population of Mardan district are given in table.8. Where the most common allele found was DYS456 (15) with the frequency of 0.720 while the least occurring alleles found in this population were DYS576 (10, 15), DYS19 (13), DYS389II (33), DYS481 (21, 27), DYS549 (15), DYS635 (20), DYS390 (20), DYS643 (15), DYS393 (15), DYS385b (20, 21) and GATAH4 (10, 14) with the frequency of 0.0133.

The most polymorphic locus in this population was DYS385b with total of 9 alleles found and the least polymorphic was DYS19 with two alleles found at this locus.

Table: 8. Allele Frequencies of 23 Y-STRs in Yousafzai (Ys) population of Mardan district.

| Alleles | DYS 576 | DYS 389I | DYS 448 | DYS 389II | DYS 19 | DYS 391 | DYS 481 | DYS 549 | DYS 533 | DYS 438 | DYS 437 | DYS 570 | DYS 635 | DYS 390 | DYS 439 | DYS 392 | DYS 643 | DYS 393 | DYS 458 | DYS 385a | DYS 385b | DYS 456 | YGATA H4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | | | | | | | | | | | | | | | | | 0.0400 | | | | | | |
| 9 | | | | | | | | | | 0.04 | | | | | | | 0.0267 | | | 0.04 | | | |
| 10 | 0.0133 | | | | | 0.56 | | | 0.0933 | 0.28 | | | | | 0.48 | 0.0667 | 0.6133 | | | | | | 0.0133 |
| 11 | | | | | | 0.44 | | 0.0933 | 0.2000 | 0.64 | | | | | 0.28 | 0.6933 | 0.1867 | 0.0667 | | 0.5867 | | | 0.1067 |
| 12 | | 0.12 | | | | | | 0.0133 | 0.6667 | 0.04 | | | | | 0.16 | 0.0400 | 0.0667 | 0.0533 | | 0.0400 | | | 0.6133 |
| 13 | | 0.64 | | | 0.0133 | | | 0.3600 | 0.0400 | | | | | | 0.08 | 0.0267 | 0.0400 | 0.7067 | | 0.2133 | 0.0400 | | 0.2267 |
| 14 | | 0.24 | | | 0.2400 | | | 0.0267 | | | 0.6800 | 0.0533 | | | | 0.1200 | | 0.1600 | 0.0400 | 0.0933 | 0.5067 | 0.0533 | 0.0133 |
| 15 | 0.0133 | | | | 0.3333 | | | 0.0133 | | | 0.1467 | 0.0933 | | | | 0.0267 | 0.0133 | 0.0133 | 0.4133 | 0.0267 | 0.0933 | 0.7200 | |
| 16 | 0.1467 | | | | 0.2533 | | | | | | 0.1733 | 0.0533 | | | | | | | 0.1867 | | 0.2000 | 0.2000 | |
| 17 | 0.2133 | | | | 0.0533 | | | | | | | 0.1600 | | | | | | | 0.2133 | | 0.0800 | 0.0267 | |
| 18 | 0.3733 | | 0.0400 | | | | | | | | | 0.2400 | 0.04 | | | | | | 0.0400 | | 0.0267 | | |
| 19 | 0.1600 | | 0.1867 | | | | 0.0667 | | | | | 0.2933 | | | | | | | 0.0800 | | 0.0267 | | |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **20** | 0.0800 | | 0.5867 | | | | | | | | | 0.1067 | 0.0133 | 0.0133 | | | | | 0.0267 | | 0.0133 | | |
| **21** | | | 0.1467 | | | | 0.0133 | | | | | | 0.0667 | 0.0267 | | | | | | | 0.0133 | | |
| **22** | | | | | | | 0.0800 | | | | | | 0.0267 | 0.0933 | | | | | | | | | |
| **23** | | | | | | | 0.4400 | | | | | | 0.5200 | 0.2400 | | | | | | | | | |
| **24** | | | | | | | 0.1867 | | | | | | 0.2133 | 0.4400 | | | | | | | | | |
| **25** | | | | | | | 0.1333 | | | | | | 0.0800 | 0.1867 | | | | | | | | | |
| **26** | | | | | | | 0.0667 | | | | | | 0.0400 | | | | | | | | | | |
| **27** | | | | | | | 0.0133 | | | | | | | | | | | | | | | | |
| **28** | | | | 0.1067 | | | | | | | | | | | | | | | | | | | |
| **29** | | | | 0.1733 | | | | | | | | | | | | | | | | | | | |
| **30** | | | | 0.4533 | | | | | | | | | | | | | | | | | | | |
| **31** | | | | 0.1467 | | | | | | | | | | | | | | | | | | | |
| **32** | | | | 0.1067 | | | | | | | | | | | | | | | | | | | |
| **33** | | | | | | | | | | | | | | | | | | | | | | | |

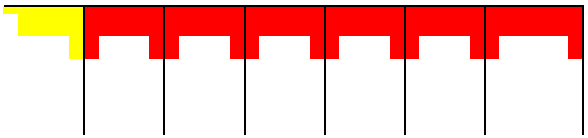
| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 0.0133 | | | | | | | | | | | | | | | | | | | |
| **GD** | **0.7715** | **0.5254** | **0.5716** | **0.7297** | **0.7151** | **0.4995** | **0.7481** | **0.6123** | **0.5121** | **0.5157** | **0.4926** | **0.8159** | **0.6782** | **0.7139** | **0.6681** | **0.4760** | **0.5768** | **0.4739** | **0.7485** | **0.6058** | **0.6941** | **0.4440** | **0.5438** |

**4.8**

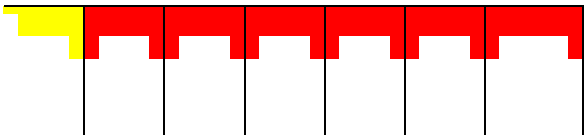
**.2 Mohmand: MD (Mardan)**

In Mohmand(MD) population of Mardan District, the most frequently occurring allele was DYS533 (12) with the highest frequency of 0.6301 while the least occurring alleles were DYS576 (21), DYS389II (27, 32), DYS19 (13), DYS549 (15), DYS635 (19, 20), DYS392 (15), DYS643 (14), DYS458 (20, 22), DYS385b (20, 22), GATAH4 (10) with the frequency of 0.0137. The most polymorphic locus among all 23 loci in this population was found DYS635 with total 9 alleles(Table: 9).

Table: 9. Allele Frequencies of 23 Y-STRs in Mohmand (MD) population of Mardan district.



| Alleles | DYS 576 | DYS 389I | DYS 448 | DYS 389II | DYS 19 | DYS 391 | DYS 481 | DYS 549 | DYS 533 | DYS 438 | DYS 437 | DYS 570 | DYS 635 | DYS 390 | DYS 439 | DYS 392 | DYS 643 | DYS 393 | DYS 458 | DYS 385a | DYS 385b | DYS 456 | YGATA H4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | | | | | | | | | | | | | | | | | | | | 0.0548 | | | |
| 8 | | | | | | | | | | | | | | | 0.0822 | | | | | | | | |
| 9 | | | | | | | | | | 0.0548 | | | | | | | 0.0685 | | | 0.1507 | | | |
| 10 | | | | | | 0.5616 | | 0.0959 | 0.0822 | 0.3973 | | | | | 0.4384 | 0.0685 | 0.5205 | | | 0.0274 | | | 0.0137 |
| 11 | | 0.0822 | | | | 0.3973 | | 0.0959 | 0.2603 | 0.5068 | | | | | 0.1918 | 0.6438 | 0.1096 | 0.0137 | | 0.4247 | | | 0.3425 |
| 12 | | 0.0548 | | | | 0.0411 | | 0.4795 | 0.6301 | 0.0411 | | | | | 0.2329 | 0.0411 | 0.2877 | 0.3562 | | 0.1644 | | | 0.5616 |
| 13 | | 0.7123 | | | 0.0137 | | | 0.2192 | 0.0274 | | | | | | 0.0548 | 0.0274 | | 0.5616 | | 0.1233 | 0.0411 | 0.0685 | 0.0822 |
| 14 | | 0.1370 | | | 0.2055 | | | 0.0959 | | | 0.5616 | | | | | 0.2055 | 0.0137 | 0.0685 | 0.0274 | 0.0274 | 0.4384 | 0.0548 | |
| 15 | 0.1233 | 0.0137 | | | 0.3699 | | | 0.0137 | | | 0.2055 | 0.0411 | | | | 0.0137 | | | 0.3425 | 0.0274 | 0.0548 | 0.5753 | |
| 16 | 0.1918 | | | | 0.3562 | | | | | | 0.2329 | 0.2466 | | | | | | | 0.1644 | | 0.1781 | 0.3014 | |
| 17 | 0.0548 | | 0.0822 | | 0.0411 | | | | | | | 0.1096 | | | | | | | 0.2055 | | 0.2055 | | |
| 18 | 0.1233 | | 0.0411 | | | | | | | | | 0.1096 | | | | | | | 0.0959 | | 0.0548 | | |
| 19 | 0.4110 | | 0.3836 | | | | 0.0685 | | | | | 0.3699 | 0.0137 | | | | | | 0.1370 | | | | |
| 20 | 0.0822 | | 0.3836 | | | | | | | | | 0.0548 | 0.0137 | | | | | | 0.0137 | | 0.0137 | | |
| 21 | 0.0137 | | 0.1096 | | | | | | | | | 0.0685 | 0.1918 | | | | | | | | | | |
| 22 | | | | | | | 0.0548 | | | | | | 0.1507 | 0.2466 | | | | | 0.0137 | | 0.0137 | | |
| 23 | | | | | | | 0.1644 | | | | | | 0.4247 | 0.2329 | | | | | | | | | |
| 24 | | | | | | | 0.2466 | | | | | | 0.1096 | 0.4521 | | | | | | | | | |
| 25 | | | | | | | 0.3288 | | | | | | 0.0274 | 0.0685 | | | | | | | | | |
| 26 | | | | 0.0822 | | | 0.0548 | | | | | | 0.0274 | | | | | | | | | | |
| 27 | | | | 0.0137 | | | 0.0822 | | | | | | 0.0411 | | | | | | | | | | |
| 28 | | | | 0.0274 | | | | | | | | | | | | | | | | | | | |
| 29 | | | | 0.2466 | | | | | | | | | | | | | | | | | | | |
| 30 | | | | 0.5205 | | | | | | | | | | | | | | | | | | | |
| 31 | | | | 0.0959 | | | | | | | | | | | | | | | | | | | |
| 32 | | | | 0.0137 | | | | | | | | | | | | | | | | | | | |
| GD | 0.7645 | 0.4703 | 0.6948 | 0.6602 | 0.6933 | 0.5323 | 0.7976 | 0.7040 | 0.5350 | 0.5887 | 0.5963 | 0.7797 | 0.7664 | 0.6853 | 0.7169 | 0.5434 | 0.6381 | 0.5605 | 0.7953 | 0.7599 | 0.7359 | 0.5784 | 0.5681 |



**4.8**

**.3 Muhammadzai: MZ (Charsada)**

Allele frequencies of all 23 Y-STRs in Muhammadzai (MZ) population are given in table: 10.

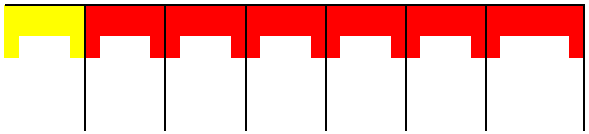
Most common alleles identified in this population were DYS392 (11), DYS456 (15) with the frequency of 0.7067. The lowest frequency of alleles found among all the 23 alleles was 0.0133 for DYS389I (10), DYS448 (18, 22), DYS391 (12), DYS481 (18, 27, 28), DYS437 (17), DYS570 (11, 14, 22), DYS635 (19), DYS390 (26), DYS439 (14), DYS392 (15), DYS643 (8, 9, 13), DYS393 (11), DYS385b (13, 21), DYS456 (13), GATAH4 (10, 14). The most polymorphic among all the loci was DYS481 with 10 alleles ranging from allele number 18 to allele numer 28.

Table: -

10. Allele Frequencies of 23 Y STRs in Muhammadzai (MZ) population of Charsada district.

| Alleles | DYS 576 | DYS 3891 | DYS 448 | DYS 3891I | DYS 19 | DYS 391 | DYS 481 | DYS 549 | DYS 533 | DYS 438 | DYS 437 | DYS 570 | DYS 635 | DYS 390 | DYS 439 | DYS 392 | DYS 643 | DYS 385a | DYS 385b | DYS 456 | YGATA H4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | | | | | | 0.04 | | | | | | | | | | | | | | | |
| 8 | | | | | | | | | | 0.0267 | | | | | | | 0.0133 | | | | |
| 9 | | | | | | 0.0267 | | | 0.0267 | 0.0800 | | | | | | | 0.1333 | | | 0.0400 | | | |
| 10 | | 0.0133 | | | | 0.4933 | | | 0.0800 | 0.2533 | | | | | 0.4800 | 0.1200 | 0.6267 | | | | | | 0.0133 |
| 11 | | | | | | 0.4267 | | 0.0267 | 0.2133 | 0.6400 | | 0.0133 | | | 0.3333 | 0.7067 | 0.1067 | 0.0133 | | 0.4933 | | | 0.4000 |
| 12 | | 0.0800 | | | | 0.0133 | | 0.6400 | 0.6533 | | | | | | 0.1333 | 0.0667 | 0.1067 | 0.2000 | | 0.0800 | | | 0.4133 |
| 13 | | 0.7600 | | | 0.0533 | | | 0.2800 | 0.0267 | | | | | | 0.0400 | 0.0267 | 0.0133 | 0.6800 | | 0.2000 | 0.0133 | 0.0133 | 0.1600 |
| 14 | 0.0533 | 0.1467 | | | 0.0533 | | | 0.0533 | | | 0.6533 | 0.0133 | | | 0.0133 | 0.0400 | | 0.1067 | 0.0267 | 0.1067 | 0.4667 | 0.0400 | 0.0133 |
| 15 | 0.0400 | | | | 0.4133 | | | | | | 0.1067 | 0.0533 | | | | 0.0133 | | | 0.4000 | 0.0400 | 0.0800 | 0.7067 | |
| 16 | 0.1733 | | | | 0.4133 | | | | | | 0.2267 | 0.0667 | | | | | | | 0.1867 | 0.0400 | 0.1067 | 0.2133 | |
| 17 | 0.0667 | | | | 0.0533 | | | | | | 0.0133 | 0.2000 | | | | | | | 0.2000 | | 0.0933 | 0.0267 | |
| 18 | 0.1867 | | 0.0133 | | | | 0.0133 | | | | | 0.1733 | | | | | | | 0.1067 | | 0.0667 | | |
| 19 | 0.3067 | | 0.2933 | | | | 0.0667 | | | | | 0.4133 | 0.0133 | | | | | | 0.0800 | | 0.0933 | | |
| 20 | 0.1467 | | 0.5733 | | | | | | | | | 0.0533 | 0.0800 | | | | | | | | 0.0667 | | |
| 21 | 0.0267 | | 0.1067 | | | | 0.0533 | | | | | 0.0133 | 0.1467 | 0.0533 | | | | | | | 0.0133 | | |
| 22 | | | 0.0133 | | | | 0.0800 | | | | | | 0.0400 | 0.1467 | | | | | | | | | |
| 23 | | | | | | | 0.3067 | | | | | | 0.4533 | 0.2533 | | | | | | | | | |
| 24 | | | | | | | 0.1200 | | | | | | 0.2000 | 0.3733 | | | | | | | | | |
| 25 | | | | | | | 0.2933 | | | | | | 0.0667 | 0.1600 | | | | | | | | | |
| 26 | | | | | | | 0.0400 | | | | | | | 0.0133 | | | | | | | | | |
| 27 | | | | | | | 0.0133 | | | | | | | | | | | | | | | | |
| 28 | | | | 0.0933 | | | 0.0133 | | | | | | | | | | | | | | | | |
| 29 | | | | 0.2267 | | | | | | | | | | | | | | | | | | | |
| 30 | | | | 0.5067 | | | | | | | | | | | | | | | | | | | |
| 31 | | | | 0.1200 | | | | | | | | | | | | | | | | | | | |
| 32 | | | | 0.0133 | | | | | | | | | | | | | | | | | | | |
| 33 | | | | | | | | | | | | | | | | | | | | | | | |
| GD | 0.8209 | 0.3996 | 0.5813 | 0.6760 | 0.6490 | 0.5798 | 0.8004 | 0.5153 | 0.5269 | 0.5279 | 0.5171 | 0.7586 | 0.7301 | 0.7564 | 0.6476 | 0.4566 | 0.5798 | 0.5106 | 0.7568 | 0.7034 | 0.7478 | 0.4587 | 0.6519 |

Table: -



**4.8**

**.4 Kakakhel Mian**

In table: 11, the allele frequencies of all the loci in Kakakhel Mian (KM) population are given. The highest allele frequency was found in this population was 0.7467 for DYS533 while the lowest allele frequency recorded in this population was 0.0133 for DYS576 (14, 15), DYS389II (20, 24, 26), DYS19 (17), DYS391 (9, 12), DYS481 (22), DYS533 (9), DYS438 (12), DYS635 (18, 20, 25, 26), DYS385a (10, 15), DYS385b (13), DYS456 (17) and GATAH4 (10).DYS635 and DYS 385b were considered to be the most polymorphic loci among all 23 loci typed in this population with 8 alleles.

Table: -

11. Allele Frequencies of 23 Y STRs in Kakakhel Mian (KM) population of Charsada district.

| Alleles | DYS 576 | DYS 389I | DYS 448 | DYS 389II | DYS 19 | DYS 391 | DYS 481 | DYS 549 | DYS 533 | DYS 438 | DYS 437 | DYS 570 | DYS 635 | DYS 390 | DYS 439 | DYS 392 | DYS 643 | DYS [illegible] | DYS 458 | DYS 385a | DYS [illegible] | DYS 456 | YGATA [illegible] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | | | | | | 0.1333 | | | 0.0133 | 0.2133 | | | | | | | 0.1467 | | | 0.0267 | | | |
| 10 | | | | | | 0.5467 | | | 0.0400 | 0.2267 | | | | | 0.4533 | 0.1200 | 0.6533 | | | 0.0133 | | | 0.0133 |
| 11 | | | | | | 0.3067 | | 0.0800 | 0.2000 | 0.5467 | | | | | 0.1867 | 0.7067 | 0.1467 | 0.1333 | | 0.3867 | | | 0.2400 |
| 12 | | 0.26667 | | | | 0.0133 | | 0.4000 | 0.7467 | 0.0133 | | | | | 0.1600 | | 0.0533 | 0.1333 | | 0.0400 | | | 0.6933 |
| 13 | | 0.65333 | | | | | | 0.5200 | | | | | | | 0.2000 | | | 0.5867 | | 0.2800 | 0.0133 | 0.0267 | 0.0533 |
| 14 | 0.0133 | 0.08000 | | | 0.2800 | | | | | | 0.5600 | | | | | 0.1467 | | 0.0533 | 0.0267 | 0.2400 | 0.4267 | 0.0267 | |
| 15 | 0.0133 | | | | 0.0800 | | | | | | 0.2533 | 0.2667 | | | | | | 0.0933 | 0.5333 | 0.0133 | 0.0400 | 0.6400 | |
| 16 | 0.2267 | | | | 0.5067 | | | | | | 0.1867 | 0.0933 | | | | 0.0267 | | | 0.1467 | | 0.0667 | 0.2933 | |
| 17 | 0.1600 | | | | 0.0133 | | | | | | | 0.1600 | | | | | | | 0.2133 | | 0.0267 | 0.0133 | |
| 18 | 0.4667 | | 0.1200 | | | | | | | | | 0.0533 | 0.0133 | | | | | | 0.0267 | | 0.0400 | | |
| 19 | 0.0933 | | 0.2533 | | | | 0.0400 | | | | | 0.2133 | | | | | | | 0.0400 | | 0.1867 | | |
| 20 | 0.0267 | | 0.5333 | 0.0133 | | | | | | | | 0.1467 | 0.0133 | | | | | | | | 0.2000 | | |
| 21 | | | 0.0933 | | | | | | | | | 0.0533 | 0.2000 | | | | | | | | | | |
| 22 | | | | | | | 0.0133 | | | | | | 0.0400 | 0.1867 | | | | | | | | | |
| 23 | | | | | | | 0.4800 | | | | | | 0.4933 | 0.2400 | | | | | | | | | |
| 24 | | | | 0.0133 | | | 0.2933 | | | | | | 0.2133 | 0.5200 | | | | | | | | | |
| 25 | | | | | | | 0.0267 | | | | | | 0.0133 | 0.0533 | | | | | | | | | |
| 26 | | | | 0.0133 | | | 0.1200 | | | | | | 0.0133 | | | | | | | | | | |
| 27 | | | | | | | | | | | | | | | | | | | | | | | |
| 28 | | | | 0.2400 | | | | | | | | | | | | | | | | | | | |
| 29 | | | | 0.1200 | | | | | | | | | | | | | | | | | | | |
| 30 | | | | 0.5200 | | | | | | | | | | | | | | | | | | | |
| 31 | | | | 0.0800 | | | | | | | | | | | | | | | | | | | |
| GD | 0.7049 | 0.50234 | 0.6368 | 0.6595 | 0.5674 | 0.5971 | 0.6572 | 0.5708 | 0.4061 | 0.6123 | 0.5953 | 0.8282 | 0.6778 | 0.6429 | 0.7034 | 0.4703 | 0.5344 | 0.6169 | 0.6446 | 0.7214 | 0.7445 | 0.5096 | 0.4649 |

**4.8**

**.5 MohmandCharsada**

Allele frequencies for all alleles are given in table: 12.The highest allele frequency was found to be 0.8243 for DYS393 in this population while the lowest allele frequency recorded was 0.0135 for DYS576(21), DYS448(18), DYS19(17), DYS549(14), DYS355(9), DYS481(18), DYS438(8), DYS570(14), DYS635(26), DYS392(10), DYS385a(16), DYS385b(20), DYS456(17) and GATAH4(14). The most polymorphic among all the loci in this population was DYS570 with 10 alleles ranging from allele number 13 to allele 21.

Table: -

12. Allele Frequencies of 23 Y STRs in Mohmand (MM) population of Charsada district.

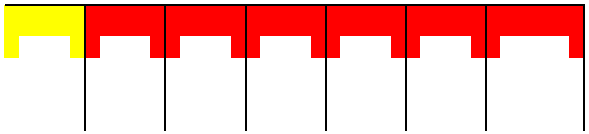
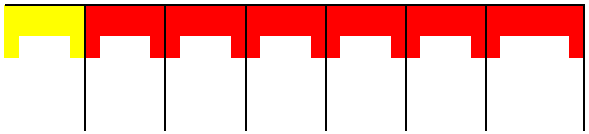


Table: -

| Alleles | DYS 576 | DYS 389I | DYS 448 | DYS 389II | DYS 19 | DYS 391 | DYS 481 | DYS 549 | DYS 533 | DYS 438 | DYS 437 | DYS 570 | DYS 635 | DYS 390 | DYS 439 | DYS 392 | DYS 643 | DYS [illegible] | DYS [illegible] | DYS [illegible] | DYS [illegible] | DYS 456 | YGATA [illegible] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | | | | | | | | | | | | | | | | | | | | | | | |
| 7 | | | | | | | | | | | | | | | | | | | | | | | |
| 8 | | | | | | | | | | 0.0135 | | | | | | | | | | | | | |
| 9 | | | | | | | | | 0.0135 | 0.0946 | | | | | | | 0.0405 | | | 0.0405 | | | |
| 10 | | | | | | 0.3784 | | | 0.1622 | 0.2297 | | | | | 0.5270 | 0.0135 | 0.7297 | | | | | | 0.0270 |
| 11 | | | | | | 0.6216 | | 0.1216 | 0.0676 | 0.5946 | | | | | 0.2568 | 0.7973 | 0.0135 | 0.0541 | | 0.6216 | | | 0.1892 |
| 12 | | 0.1216 | | | | | | 0.4459 | 0.7027 | 0.0676 | | | | | 0.1081 | | 0.2162 | 0.0811 | | 0.0405 | | | 0.5541 |
| 13 | | 0.7973 | | | 0.0270 | | | 0.4189 | 0.0541 | | | 0.0270 | | | 0.0946 | 0.0405 | | 0.8243 | | 0.2162 | | 0.0405 | 0.2162 |
| 14 | | 0.0811 | | | 0.1757 | | | 0.0135 | | | 0.5811 | 0.0135 | | | 0.0135 | 0.0811 | | 0.0405 | | 0.0676 | 0.5811 | 0.0270 | 0.0135 |
| 15 | 0.0676 | | | | 0.2568 | | | | | | 0.2162 | 0.0270 | | | | 0.0270 | | | 0.4189 | | 0.0811 | 0.5946 | |
| 16 | 0.2162 | | | | 0.5270 | | | | | | 0.2027 | 0.1081 | | | | 0.0405 | | | 0.2432 | 0.0135 | 0.2297 | 0.3243 | |
| 17 | 0.1486 | | | | 0.0135 | | | | | | | 0.2432 | | | | | | | 0.1351 | | 0.0405 | 0.0135 | |
| 18 | 0.1892 | | 0.0135 | | | | 0.0135 | | | | | 0.1081 | | | | | | | 0.0811 | | 0.0541 | | |
| 19 | 0.2838 | | 0.2432 | | | | 0.1351 | | | | | 0.3514 | | | | | | | 0.0405 | | | | |
| 20 | 0.0811 | | 0.5541 | | | | | | | | | 0.0676 | | | | | | | 0.0541 | | 0.0135 | | |
| 21 | 0.0135 | | 0.1892 | | | | 0.0405 | | | | | 0.0541 | 0.0676 | 0.0405 | | | | | | | | | |
| 22 | | | | | | | 0.0405 | | | | | | 0.0405 | 0.0811 | | | | | 0.0270 | | | | |
| 23 | | | | | | | 0.3919 | | | | | | 0.4595 | 0.2027 | | | | | | | | | |
| 24 | | | | | | | 0.1216 | | | | | | 0.4054 | 0.5405 | | | | | | | | | |
| 25 | | | | | | | 0.2162 | | | | | | 0.0135 | 0.1351 | | | | | | | | | |
| 26 | | | | | | | 0.0405 | | | | | | 0.0135 | | | | | | | | | | |
| 27 | | | | | | | | | | | | | | | | | | | | | | | |
| 28 | | | | 0.081 | | | | | | | | | | | | | | | | | | | |
| 29 | | | | 0.243 | | | | | | | | | | | | | | | | | | | |
| 30 | | | | 0.527 | | | | | | | | | | | | | | | | | | | |
| 31 | | | | 0.135 | | | | | | | | | | | | | | | | | | | |
| 32 | | | | 0.014 | | | | | | | | | | | | | | | | | | | |
| GD | 0.8145 | 0.3477 | 0.6061 | 0.647 | 0.6331 | 0.4769 | 0.7719 | 0.6220 | 0.4787 | 0.5879 | 0.5812 | 0.7956 | 0.6264 | 0.6490 | 0.6442 | 0.3584 | 0.4247 | 0.3136 | 0.7453 | 0.5665 | 0.6064 | 0.5461 | 0.6179 |

Table: -



### 4.9 Principle Coordinate Analysis

A Principle Coordinate (PCoA) plots was constructed using data from the five populations from the study area and other neighboring populations available on YHRDwith accession number (Afghanistan (Pathan): YC000226; Xinjiang (Uighur): YA004122;Tibet China(Tibetan): YA004005; Xingjiang(Kazakh): YA003848; Greece(Greeks): YC000125; India (Tamil): YC000123; Israel: YC000064;Iran: YA003782; Iraq (Iraqis): YC000007; Russia: YC000079; Turkey: YC000296, YC000173, YA003467; Pakistan (Yousafzai): YC000226; Pakistani (Pathan): YC000147).The clustering pattern revealed that all the Pathan populations (including previously studied Pathan populations) from Pakistan and Afghanistan are clustered together which shows great homogeneity with each other and little genetic sharing with Russian populations as well (Fig.23).

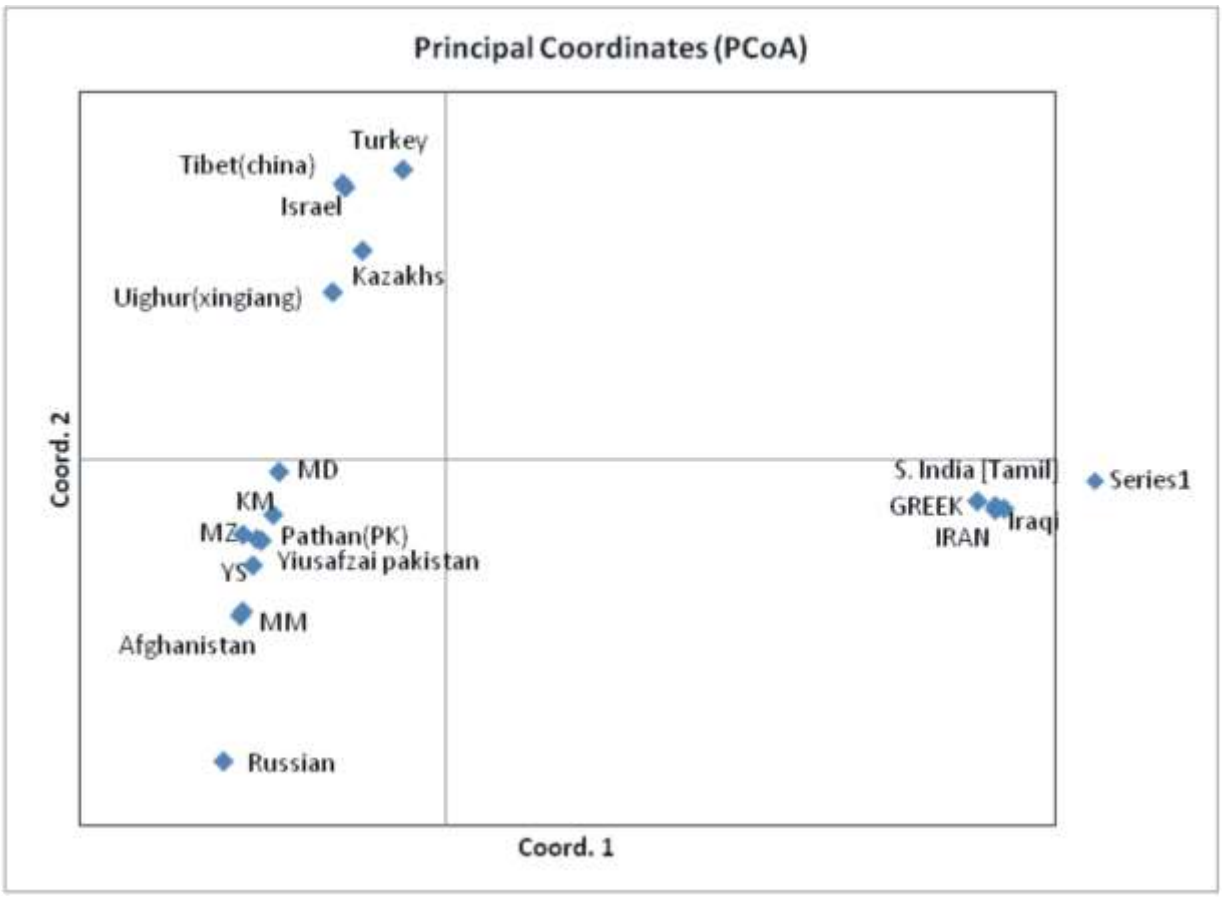



Fig.23. Y-$F_{ST}$ based PCoA plot of five populations of Charsada and Mardan and comparative global populations *(The abbreviations MZ, KM, MM, YS, MD denote the tribes Muhammadzai, Kakakhel Mian, Mohmand from Charsada, Yousafzai and Mohmand from Mardan, respectively).*

**4.10 mtDNA HVI analysis**

A total of 93 haplotypes were identified in 165 samples (table: 15). 38.7% of the haplotypes were found to be shared among all the populations and 57 haplotypes (61.3%) were singletons.

### 4.10.1 Genetic Diversity

Genetic diversity in all the samples amplified for HVI control region was 0.993 and

Match probability was 0.987 (table. 13).

Table 13. Comparison of *mt*DNA Genetic diversity among various ethnic groups of Pakistan.

| Parameters | **Path** (this study) | **Mkr** (Siddiqi *et al.* 2015) | **Pth** (Rakha *et al.*2011) | **Bal** (Murci *et al.* 2004) | **Brh** (Murci *et al.* 2004) | **Haz** (Murci *et al.* 2004) | **Bsh** (Murci *et al.* 2004) | **Ksh** (Murci *et al.* 2004) | **Par** (Murci *et al.* 2004) | **Snd** (Murci *et al.* 2004) | **Sar** (Hayat *et al.* 2015 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| No. of Sample | 165 | 100 | 230 | 39 | 38 | 23 | 44 | 44 | 44 | 23 | 85 |
| No. of haplotypes | 93 | 70 | 157 | 26 | 22 | 21 | 32 | 12 | 22 | 21 | 63 |
| No of unique haplotypes | 57 | 54 | 128 | 18 | 15 | 19 | 25 | 5 | 12 | 19 | 58 |
| Genetic Diversity | 0.993 | 0.968 | 0.993 | 0.974 | 0.952 | 0.992 | 0.980 | 0.851 | 0.950 | 0.992 | 0.957 |

*Note: The abbreviations Path, Mkr, Pth, Bal, Brh, Haz, Bsh, Ksh, Par, Snd and Sar stands for Pathan, Makrani, Pathan, Baluch, Brahui, Hazara, Burusho, Kalash, Parsi, Sindhi and Saraiki respectively*

### 4.10.2 mtDNA statistics

Types and frequency of variations observed in the mitochondrial HVI sequences of all the five populations is given in the table: 14.

Table: 14. Statistic parameters of five populations of Charsada and Mardan.

| **Population** | **MZ(n=28)** | **KM(n=35)** | **MM(n=31)** | **YS(n=36)** | **MD(35)** | **Mean** | **S.d** |
|---|---|---|---|---|---|---|---|
| **No. of Transitions** | 22 | 18 | 35 | 15 | 20 | 22.0 | 7.714 |
| **No. of Transversion** | 4 | 6 | 3 | 2 | 4 | 3.8 | 1.483 |
| **No. of Substitutions** | 26 | 24 | 38 | 17 | 24 | 25.8 | 7.629 |
| **No. of Indels** | 0 | 2 | 4 | 1 | 2 | 1.8 | 1.483 |

*Note: The abbreviations MZ, KM, MM, YS, MD and S.d stands for Muhammadzai, Kakakhel Mian, Mohmand from Charsada, Yousafzai, Mohmand from Mardan and Standard deviation respectively.*

Table: 15.List of SNPs defining mtDNA HVI Haplogroups in five Pathan populations of Charsada and Mardan district.

| Samples | Macro Haplogroup | Haplogroup | Frequency | Origin | No.of variants | SNPs |
|---|---|---|---|---|---|---|
| KM01 | R | U2e | 4 | S.Asian | 7 | A16051G, G16129C, A16182C, A16183C, T16189C, A16194C, C16197G |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| KM02 | R | H2(T152C T16311C) | | SW.Asian/Middle Eastern | 5 | 1 | T16311C |
| KM04 | M | M3 | | S.Asian | 7 | 4 | T16126C, C16223T, T16311C, T16519C |
| KM06 | R | J1b1a1 | | Middle eastern | 5 | 7 | C16069T, T16126C, G16145A, T16172C, C16222T, C16261T, G16474T |
| KM08 | R | U5a2 | | S.Asian | 1 | 5 | C16192T, C16256T, C16270T, C16278T, G16526A |
| KM10 | R | U2c | | SW.Asian/Middle Eastern | 2 | 2 | A16051G, C16179T |
| KM11 | R | I1 | | SW.Asian/Middle Eastern | 1 | 8 | G16129A, A16180G, C16223T, A16254G, A16293C, T16311C, G16391A, T16519C |
| KM12 | R | R6(G16129A) | | S.Asian | 3 | 4 | G16129A, G16213A, T16362C, T16519C |
| KM13 | R | J1c | | S.Asian | 1 | 6 | C16069T, T16126C, C16168T, C16266T, C16278T, T16311C |
| KM14 | R | U6c | | SW.Asian/Middle Eastern | 3 | 5 | G16129A, C16169T, A16183C, T16189C, A16194C |
| KM15 | M | M5a1 | | Middle eastern | 1 | 5 | G16129A, C16223T, C16234T, C16291T, T16519C |
| KM16 | R | H1 | | S.Asian | 1 | 2 | A16051G, C16239T |
| KM17 | R | J1b5b | | S.Asian | 2 | 7 | C16069T, T16126C, G16145A, C16261T, T16263C, C16290T, T16519C |
| KM18 | R | U2e1f | | SW.Asian/Middle Eastern | 1 | 11 | A16051G, G16129C, A16182C, A16183C, T16189C, A16194T, C16221T, T16311C, T16325C, T16362C, C16447T |
| KM20 | M | M33 | | S.Asian | 1 | 4 | C16223T, T16324C, T16357C, C16527T |
| KM25 | M | M6a1b | | SW.Asian/Middle Eastern | 2 | 5 | C16188T, C16223T, T16231C, T16362C, T16519C |
| KM27 | R | R8a1a3 | | C.Asian | 2 | 4 | C16259T, C16292T, A16497G, T16519C |
| KM28 | M | M4 | | S.Asian | 3 | 7 | T16086C, C16111T, G16145A, C16223T, C16261T, T16311C, T16519C |
| KM30 | R | H6 | | S.Asian | 1 | 2 | T16362C, A16482G |
| KM31 | N | N7 | | SW.Asian/Middle Eastern | 2 | 3 | G16129A, C16223T, T16519C |
| KM32 | R | U2e1h | | SW.Asian/Middle Eastern | 2 | 6 | A16051G, G16129C, A16183C, C16193CC, T16362C, T16519C |
| KM34 | M | C4a | | Middle eastern | 1 | 5 | T16093C, G16129A, T16298C, C16327T, T16519C |
| KM35 | R | R2 | | S.Asian | 5 | 2 | C16071T, T16519C |
| KM38 | R | H10((T16093C)) | | SW.Asian/Middle Eastern | 3 | 4 | T16093C, C16278T, G16319A, T16519C |
| MD02 | R | H2a2a1d | | 2 | 3 | | T16172C, A16182AC, A16183C |
| MD03 | N | N1a1a1a | | 1 | 7 | | C16147A, T16172C, C16223T, C16248T, C16320T, C16355T, T16519C |
| MD04 | R | U2a | | 1 | 3 | | A16051G, A16206C, T16311C |
| MD05 | R | H2a3 | | 1 | 5 | | T16093C, C16223T, C16234T, G16274A, T16519C |
| MD07 | N | W4 | 2 | SW.Asian/Middle Eastern<br>E.Asian<br>IndoEuropian | 5 | | T16086C, C16223T, C16286T, C16292T, T16519C |
| MD08 | N | W | 2 | IndoEuropian | 5 | | C16192T, C16223T, A16284G, C16292T, T16519C |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| MD09 | R | H2a2a1g | 6 | 4 | T16172C, A16182AC, A16183C, T16189C | MD10 | M | G2a1d2 |
| 1 | 4 | C16223T, A16227G, C16278T, T16362C | MD12 | R | U4a3 | 1 | 3 | T16356C, T16362C, T16519C |
| MD13 | R | U7 | | | S.Asian | 4 | 5 | A16309G, A16318T, A16343G, T16362C, T16519C |
| MD18 | R | U5a1b(T16362C) | | | S.Asian | 2 | 6 | C16192T, C16256T, C16270T, T16362C, A16399G, T16519C |
| MD20 | M | M5 | | | S.Asian | 2 | 3 | G16129A, C16223T, T16519C |
| MD23 | M | M65a(C16311T) | | | S.Asian | 2 | 5 | T16126C, C16223T, A16289G, A16399G, T16519C |
| MD25 | M | M5a2a1a | | | S.Asian | 1 | 4 | G16129A, C16223T, A16265C, T16519C |
| MD26 | R | H94 | | | SW.Asian/Middle Eastern | 1 | 2 | C16339T, C16355T |
| MD29 | N | N | | | C.Asian | 6 | 3 | C16223T, C16292T, T16519C |
| MD31 | R | H5 | | | SW.Asian/Middle Eastern | 4 | 3 | C16266T, T16304C, T16519C |
| MD34 | M | D4q | | | E.Asian | 1 | 6 | C16223T, C16256T, T16311C, G16319A, T16362C, T16519C |
| MD36 | M | M52a | | | S.Asian | 1 | 5 | C16223T, A16275G, G16303A, G16390A, T16519C |
| MD38 | M | M71 | | | S.Asian | 1 | 2 | C16223T, T16271C |
| MD39 | R | T1a | | | W.Eurasian | 1 | 6 | T16126C, A16163G, C16186T, T16189C, C16294T, T16519C |
| MM01 | R | P6 | | | SE.Asian | 1 | 6 | C16221T, T16311C, T16325C, T16362C, C16447T, T16519C |
| MM02 | M | M2a1 | | | S.Asian | 1 | 7 | T16075C, C16223T, C16270T, G16274A, G16319A, T16352C, T16519C |
| MM03 | R | H15a | | | SW.Asian/Middle Eastern | 2 | 1 | C16184T |
| MM07 | R | T1a1'3 | | | W.Eurasian | 2 | 6 | T16126C, A16163G, C16186T, T16189d, C16294T, T16519C |
| MM11 | M | M49 | | | S.Asian | 2 | 3 | C16223T, C16234T, T16519C |
| MM13 | R | U2e3 | | | S.Asian | 1 | 4 | C16168T, C16234T, C16260T, T16362C |
| MM15 | R | H13a1d | | | SW.Asian/Middle Eastern | 1 | 1 | C16234T |
| MM20 | N | Y2 | | | C.Asian | 1 | 5 | T16126C, C16223T, C16266T, T16311C, T16519C |
| MM22 | R | H2a2b | | | SW.Asian/Middle Eastern | 1 | 2 | C16291T, T16311C |
| MM23 | R | H1e | | | S.Asian | 2 | 3 | G16129A, A16182AC, A16183C |
| MM26 | R | R0a | | | SW.Asian/Middle Eastern | 2 | 3 | T16126C, T16362C, T16519C |
| MM27 | R | H4a | | | S.Asian | 1 | 1 | C16287T |
| MM29 | M | M18b | | | SE.Asian | 1 | 5 | A16160G, C16223T, A16318T, T16325C, T16519C |
| MM30 | R | P4a | | | SW.Asian/Middle Eastern | 1 | 3 | A16037G, C16111T, G16319A |
| MM35 | R | H14b1 | | | C.Asian | 1 | 2 | T16126C, T16519C |
| MM36 | M | C4a1(G16129A) | | | S.Asian | 4 | 5 | G16129A, C16223T, T16298C, C16327T, T16519C |
| MM38 | M | M2b | | | SW.Asian/Middle Eastern | 1 | 1 | C16169CC |
| Mz04 | R | H1bt | | | SE.Asian | 1 | 6 | C16292T, C16355T, T16406d, A16497G, T16519C, C16527T |
| MZ05 | R | F1c1a | | | | 1 | 3 | C16111T, G16129A, T16304C |
| MZ06 | N | W6 | 2 | | IndoEuropian | 6 | | C16192T, C16223T, C16292T, T16325C, C16465T, T16519C |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| MZ11 | R | H13a1a1d | | SW.Asian/Middle Eastern | | | | 1 6 | T16086C, G16145A, C16173T, C16261T, T16311C, T16519C |
| MZ16 | R | I6a | | SW.Asian/Middle Eastern | | | | 1 8 | G16129A, C16223T, A16293C, T16311C, T16362C, G16391A, C16447T, T16519C |
| MZ22 | R | H3p | | SW.Asian/Middle Eastern | | | | 1 3 | C16069G, C16222T, A16269G |
| MZ24 | R | J1b3 | | Middle eastern | | | | 1 6 | C16069T, T16126C, G16145A, C16222T, A16235G, C16261T |
| MZ27 | M | M39a | | S.Asian | | | 1 1 | C16353T | |
| MZ28 | R | T1 | | W.Eurasian | | | | 1 7 | T16093C, T16126C, A16163G, C16186T, T16189C, C16294T, T16519C |
| MZ33 | R | H14a | | SW.Asian/Middle Eastern | | | | 1 3 | C16256T, C16270T, T16352C |
| MZ34 | R | T2e | | W.Eurasian | | | | 1 5 | T16126C, G16153A, C16294T, C16296T, T16519C |
| MZ36 | R | U2b2 | | S.Asian | | | 1 2 | A16051G, C16239T | |
| YS01 | M | M2a1a | | S.Asian | | | | 2 6 | C16223T, C16270T, G16319A, T16352C, C16449T, C16451T |
| YS03 | R | H2a2a1c | | SW.Asian/Middle Eastern | | | | 4 8 | A16051G, T16086C, C16259A, C16267T, C16291T, A16300G, A16326G, C16353T |
| YS04 | R | JT | | Middle eastern | | | | 1 3 | T16086C, T16126C, T16519C |
| YS06 | R | HV2 | | W.Eurasian | | | 1 2 | T16217C, C16446T | |
| YS07 | R | U1a2 | | S.Asian | | | | 1 4 | G16129A, T16189C, C16192T, A16202C |
| YS09 | R | J1b4a | | Middle eastern | | | | 1 6 | C16069T, T16126C, G16145A, C16218T, C16261T, C16287T |
| YS11 | R | H17a1 | | SW.Asian/Middle Eastern | | | | 1 4 | G16129A, C16223T, C16291T, T16519C |
| YS12 | M | C4a2'3'4 | | C.Asian | | | | 1 6 | C16223T, T16297C, T16298C, C16327T, T16357C, T16519C |
| YS13 | R | U2a1a | | S.Asian | | | | 1 6 | A16051G, T16154C, A16206C, A16230G, T16311C, T16519C |
| YS19 | R | T2b2b | | C.Asian | | | | 1 5 | C16111A, T16126C, C16294T, C16296T, T16519C |
| YS20 | M | M3c2 | | SW.Asian/Middle Eastern | | | | 1 4 | T16126C, T16154C, C16223T, T16519C |
| YS21 | M | C4a4b | | Middle eastern | | | | 1 8 | T16086C, G16129A, C16150T, C16223T, T16298C, C16327T, T16357C, T16519C |
| YS23 | R | H1k | | S.Asian | | | | 1 4 | A16051G, T16189C, C16290T, C16292T |
| YS24 | R | J1b | | C.Asian | | | 2 5 | C16069T, T16126C, G16145A, C16261T, T16519C | |
| YS30 | R | R | | | | | 1 3 | C16292T, A16497G, T16519C | |
| YS31 | N | N1b1 | | | | | 1 7 | G16145A, C16176G, C16223T, C16256T, A16309G, G16390A, T16519C | |
| YS36 | L3 | L3 | 1 | | E.African | | | 5 | T16093C, G16129A, C16223T, A16305T, T16519C |
| YS37 | M | M38d | | S.Asian | | | | 1 5 | G16129A, C16223T, C16266T, T16311C, T16519C |
| YS38 | R | R6b | | E.Asian | | | | 2 7 | C16179T, A16227G, C16245T, C16266T, G16274A, C16278T, T16362C |

### 4.10.3 Mitochondrial DNA HVI Haplogroups found

To characterize the maternal genetic variation among the populations, haplogroup frequencies were calculated. Among the 93 haplotypes, 63.4% of the samples belong to the haplogroup R while haplogroup M, N and L were found with the frequency of 26.8%, 8.6% and 1.1% respectively (Fig.24).

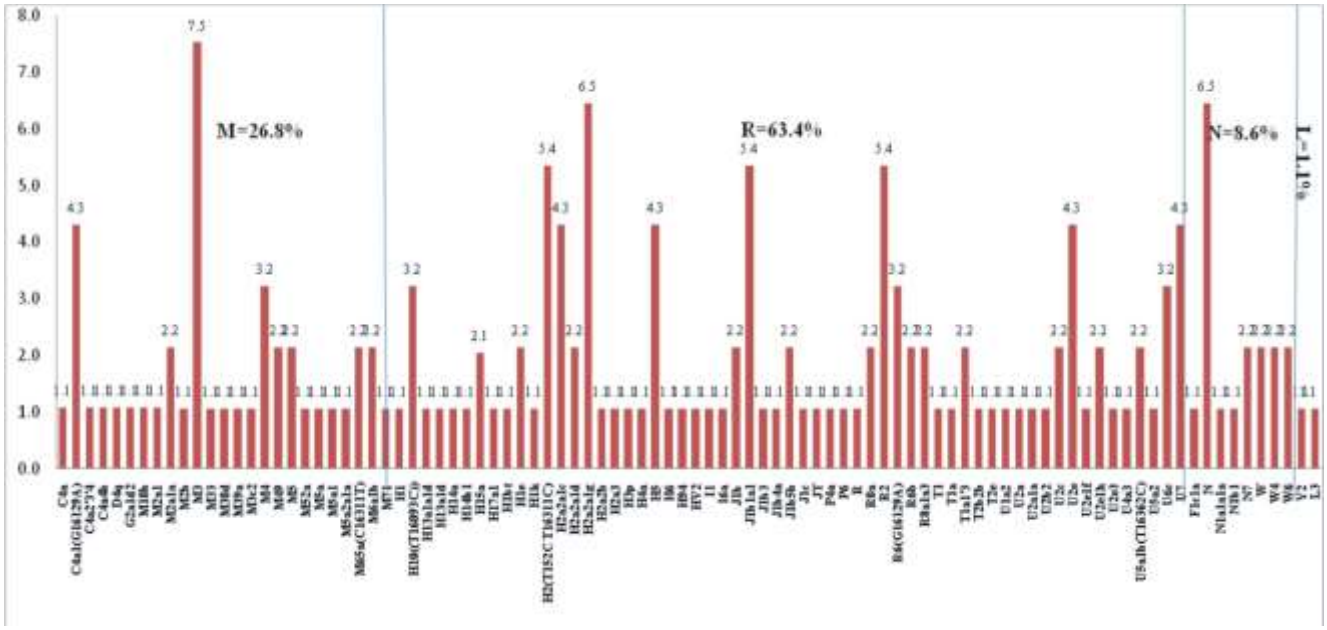



Fig.24. mtDNA HVI Haplogroup frequencies among five Pathan populations of Charsada and Mardan.

### 4.10.4 Network Analysis

To obtain insight into the genetic structure of the studied populations Median joining Networks were constructed for all the major haplogroups based on the frequency of haplotypes in these five populations using program NETWORK 4.6.1.3.

### 3.10.5 Macro Haplogroup R

Little sharing of haplotypes was observed among populations in haplogroup R and a "star haplotype" with total frequency of 47.8% in MM, 17.4% in MD, 13.04% in MZ and KM and 8.7% in YS was found in this haplogroup sequences (Fig.25a).

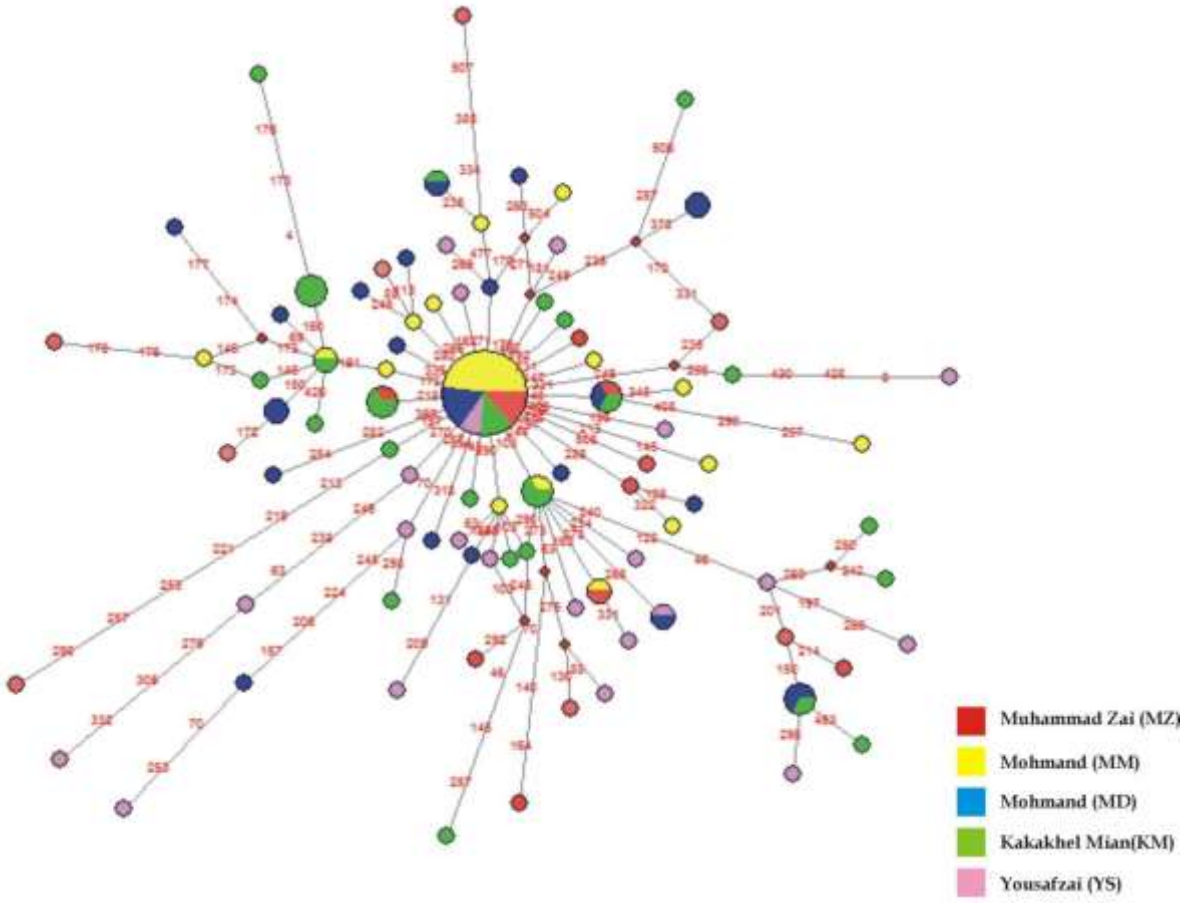



Fig.25.a. Median-Joining network of mtDNA haplogroup R in five populations of Charsada and Mardan. Areas of circles are proportional to the haplotype frequencies.

### 4.10.6 Macro Haplogroup M

Very little sharing of haplotypes was observed in Hg M. and (Fig.25 b).

There was no haplotype found shared among all the five populations. Two haplotypes were shared between Mohmand from Charsada (MM) and Yousafzai (Ys) from Mardan.Similarly single haplotype was shared between Kakakhel Mian (KM) and Muhammadzai (MZ).

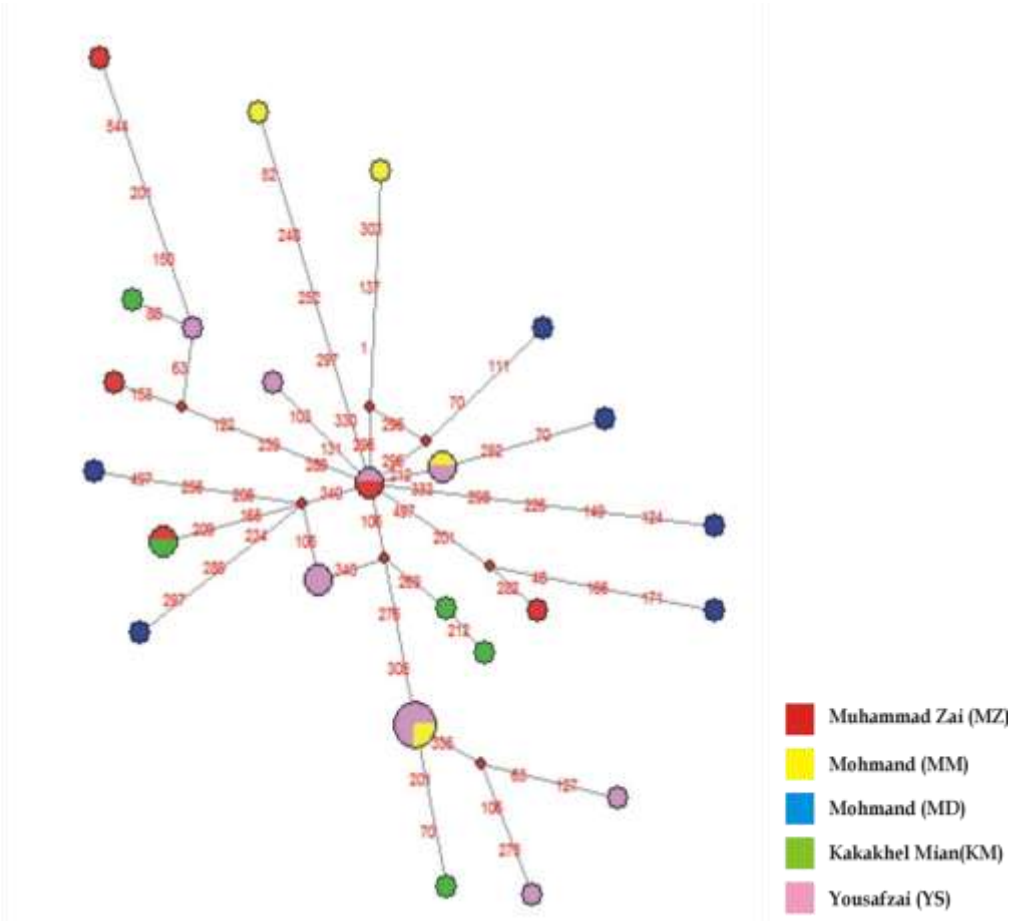



Fig.25.b. Median-Joining network of mtDNA haplogroup M in five populations of Charsada and Mardan. Areas of circles are proportional to the haplotype frequencies.

### 4.10.7 Macro Haplogroup N

This haplogroup is present in 8.6% of the samples and no haplotype sharing was recorded in Hg N (Fig. 25c).

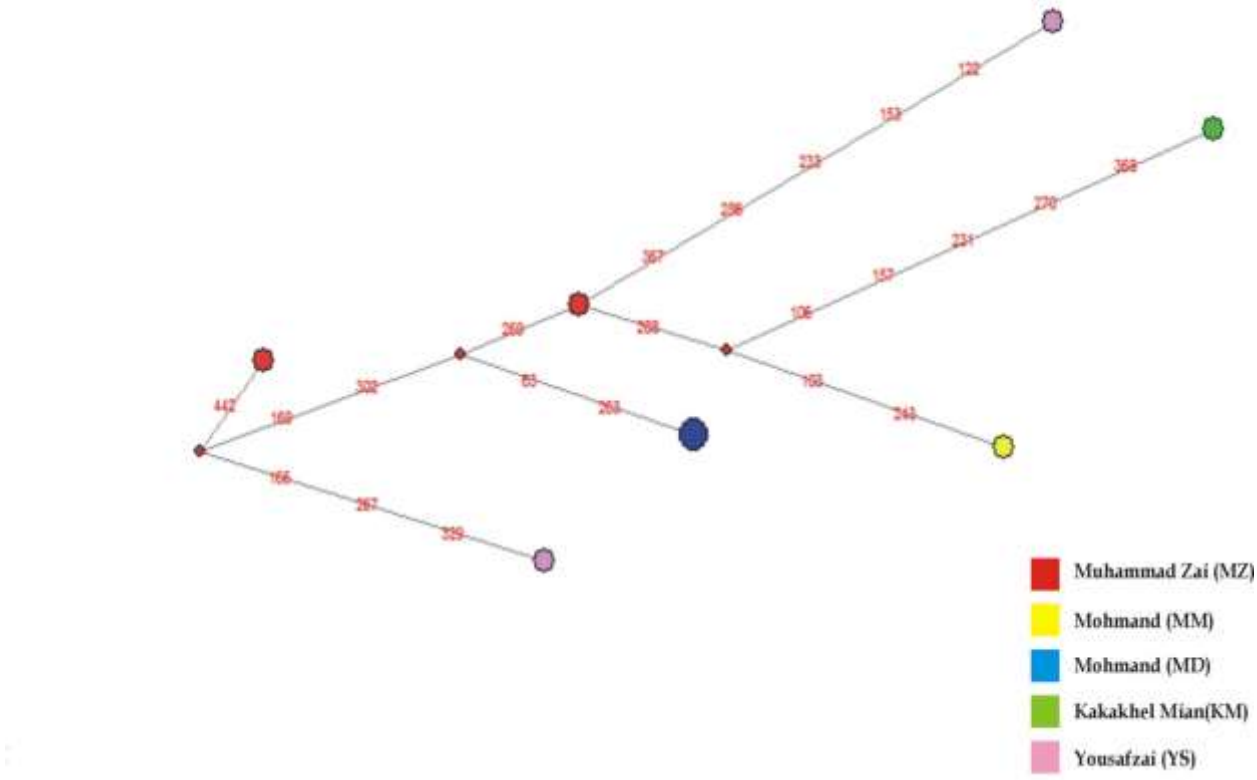



Fig.25.c. Median-Joining network of mtDNA haplogroup N in five populations of Charsada and Mardan. Areas of circles are proportional to the haplotype frequencies. mtDNA HVI $F_{ST}$ among all the five populations is represented in the fig.26.

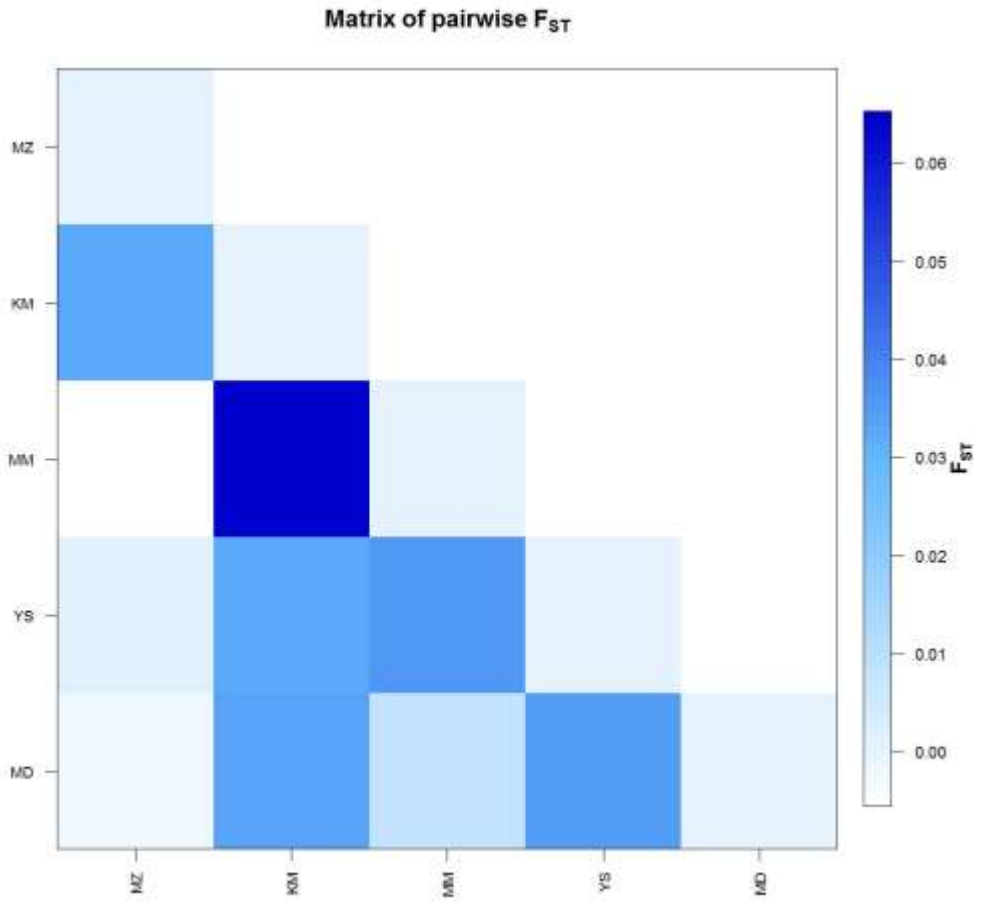



Fig.26. mtDNA HVI based Pairwise genetic distance ($F_{ST}$) among five population pairs. *(The abbreviations MZ, KM, MM, YS, MD denote the tribes Muhammadzai, Kakakhel Mian, Mohmand from Charsada, Yousafzai and Mohmand from Mardan, respectively).*

**4.10.8 Analysis of Molecular variance (AMOVA)**

The proportion of genetic variation distributed within and between the populations of two districts was assessed by analysis of molecular variance (AMOVA). Percentage of variation among groups was -1.57% and variation among populations within the group was 3.63% while percentage of variation within populations was found to be 97.94%

(table: 16).

Table: 16. AMOVA Analysis (mtDNA HVI Haplotypes) of five populations from two districts (groups).

| Source of Variation | Degree of freedom | Sum of Squares | Variance Components | Percentage of Variation |
|---|---|---|---|---|
| **Among Groups** | 1 | 1.270 | -0.02156 Va | -1.57 |
| **Among population within Groups** | 3 | 9.600 | 0.04977 Vb | 3.63 |
| **Within populations** | 182 | 244.445 | 1.34311 Vc | 97.94 |
| **Total** | 186 | 255.316 | 1.37131 | |

### 4.10.9 Principle Coordinate Analysis

$F_{ST}$ values were calculated for all pair of populations in Arlequin V.3.5.2.2 (Excoffier *et al*., 2010). No significant genetic distance ($F_{ST}$) was observed between population pairs except for the Kakakhel Mian and Mohmand. The PCoA plot provided information regarding the degree of association between different indigenous and global populations using comparative sequence data from Greece, Russia, Iraq, Iran, Israel, Afghanistan, Xinjiang china, Turkey, South India, North Indian Afridi, Balouchi, Brahui, Brusho, Kalash and Pakistani Pathan population. Clustering pattern revealed that Mohmand (MD) and Yousafzai from Mardan are clustered together with Balouchi and Pathans from Pakistan and Afridi pathans from India while Muhammadzai and Kakakhel Mian from Charsada are clustered together with Brahui and Brusho from Pakistan. On the other hand Mohmand (MM) from Charsada clusters with Turkish population,whereas Tamil from

South India and Kalash from Pakistan clustered separately (Fig.27).

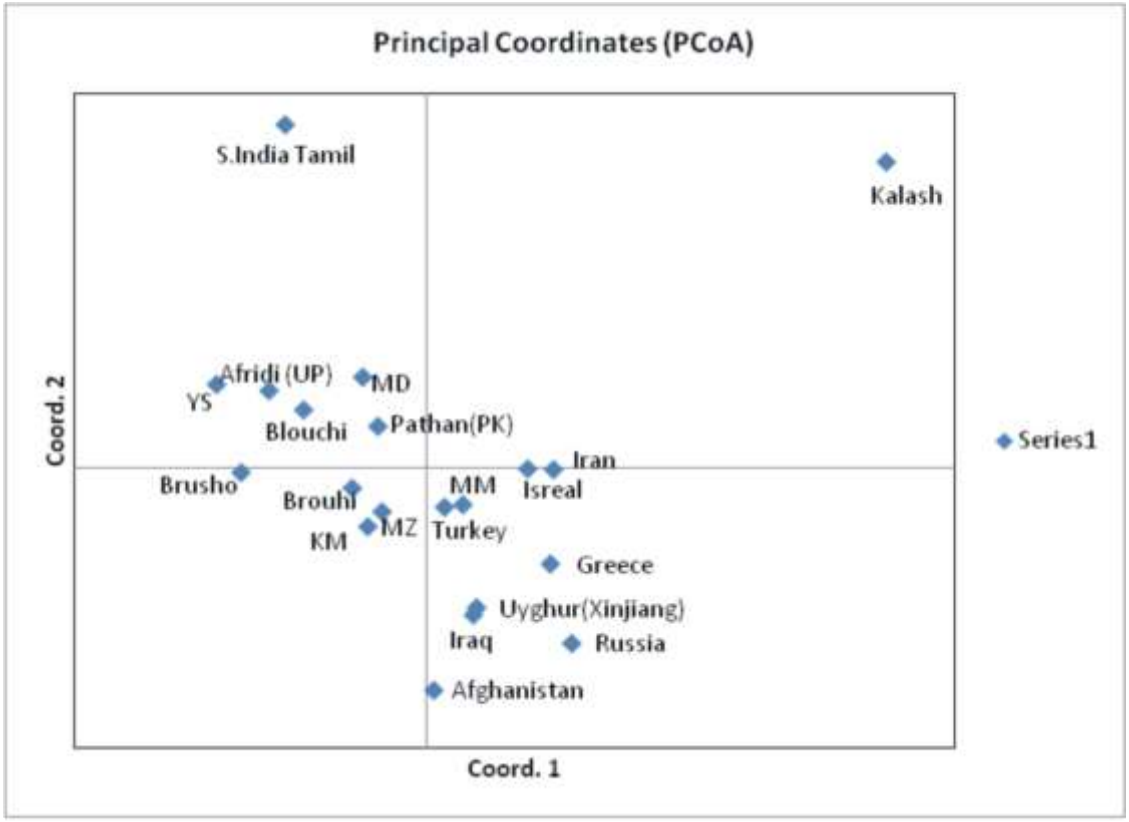



Fig.27. mtDNA-$F_{ST}$ based PCoA plot of five populations of Charsada and Mardan and comparative indigenous Pakistani and global populations. *(The abbreviations MZ, KM, MM, YS, MD, denote the tribes Muhammadzai, Kakakhel Mian, Mohmand from Charsada, Yousafzai and Mohmand from Mardan, respectively).*

**4.11 Spatial correlation of data**

Result of the Mantel Test revealed that no significant correlation existed between the genetic distance obtained from Y-Fst or *mt*DNA-Fst and geographic distance. A very little positive correlation (Table: 17) was recorded between Y-Fst and mtDNA-Fst values.

Table17. Correlation between genetic distance based on Y-Fst or mtDNA-Fst and geographic distance

| Y Chromosome, mtDNA and Geography | | Coefficient | P value |
|---|---|---|---|
| Correlation Coefficient | Y-chromosome / mtDNA | 0.21 | 0.13 |

| | | | |
|---|---|---|---|
| | Y-chromosome / Geography | -0.45 | 0.04 |
| | mtDNA / Geography | -0.07 | 0.46 |

**Chapter-5**

# DISCUSSION

Several Pathan tribes are residing in Pakistan, whose origin and affinities are still question mark. A research endeavor was carried out to elaborate Pathan population of Mardan and Charsada using 17 STR loci with high Discrimination Capacity reported by Lee *et al*. (2014). It reported that the 17 loci in Y-filer had low haplotype diversity in Pathan populations of Pakistan, but diversity was considerably increased with additional useful STR loci. In the present study, increased haplotype diversity and considerably increased discrimination capacity was observed in different Pathan populations of Charsada and Mardan with the use of 23 Y-STR loci which signals that the additional six loci in Promega PowerPlex® Y23 System are very useful for genetic investigations of these populations. Pathans generally have been reported to have low haplotype diversity due to patrilocality (Vermeulen, M. *et al*., 2009; Mohyuddin, 2001; Ghosh, 2011) but here the increased haplotype diversity has been observed among these tribes using 23 Y-STR loci. The most polymorphic locus among them was DYS570 and being rapidly mutating (RM) locus such markers are good candidates for population differentiation (Ballantyne *et al*., 2010; Ballantyne *et al*., 2012).

Genetic finding of Haber *et al*., (2012) shows that the Pathans of Afghanistan exhibit genetic affinity with the North and west India and they split from rest of the afghans 4.7kya during the rise of area" s first civilizations at Indus valley and Bactria-Margiana. The main lineage in Pathans of Afghanistan has the oldest coalescent time of 14kya in Indus valley.

The Y-haplogroup frequencies among the Pathan populations of Mardan and Charsada Districts show that they exhibit a blend of haplogroups of different origins, with contribution of 71.3% of South Asian, 15.73% from Middle Eastern, Caucasian or GrecoAnatolian and 10.4% of Central Asian origin besides little or negligible percentage of haplogroups of African origin, e.g. haplogroup E1b1a and E1b1b which are sub Saharan African and Arabian specific haplogroups. The E1b1b1-M35 lineages in some Pakistani Pathan were previously traced to a Greek origin brought by Alexander" s invasions (Firasat *et al*., 2007). However, RM network analysis of E1b1b1-M35 revealed that Afghanistan" s lineages were correlated with Middle Easterners and Iranians but not with populations from the Balkans. The Islamic invasion in the 7th century CE left an immense cultural impact on the region, with reports of Arabs settling in Afghanistan and mixing with the local population (Emadi, 2005). However the genetic signal of this expansion is not clearly evident as some Middle Eastern lineages such as E1b1b1-M35 were also found in Afghanistan.

The high frequency of Y haplogroup R1a in all of the five populations has been noticed. Almost equal frequency of this haplogroup in Pathans from Afghanistan has been

observed that indicates the genetic similarity between the two (Lacau *et al*., 2012) and similarly equal frequency of this haplogroup has been recorded in Pakistani Pathans also (Lee *et al*., 2014). South Central Asia i.e. Pakistan and India (Underhill *et al*., 2010; Mirbal *et al*., 2009) are proposed as the most likely places of origin because in both regions R1a1a-M198 has been observed at frequencies 50% or above. According to Underhill *et al*, 2010, the highest Y-STR haplotype diversity for the R1a1a lineage is observed in South Central Asia with a coalescent time of 14kya, suggesting that this region is the likely source of the dispersal for the M198 mutation. Alternatively, Klyosov claims that haplogroup R1a1a-M198 originated in South Siberia about 20kya. But later the South Central Asian origin of this haplogroup was well supported.Our results revealed the frequency of this haplogroup ranging from 41.3% to 54.6% among all the five populations of this region providing a clear evidence for the local origin of this haplogroup.

Similarly, Y haplogroup L being second most prevalent haplogroup comprising 17.9% of the total haplotypes, has already been reported in Pakistani Pathans with frequency of approximately 7% and with moderate frequency of approximately 2% among other Pakistani populations, whereas its high frequency approximately 23% was reported in Pathans of Afghanistan (Firasat *et al*., 2007). The Y chromosomal DNA variation in Pakistan by Qamar *et al*., (2002) reported 14% frequency of Haplogroup L where it was called haplogroup 28, in overall Pakistani population (Qamar *et al*., 2002), while 5.9% of this haplogroup had been reported in Pakistani Pathans in 2014 (Lee *et al*., 2014), which is said to be related with the arrival of agriculture in this area of Indian-subcontinent (Qamar *et*

*al*., 2002). Haplogroup L-M20, was hypothesized to have originated in India or the Middle East (Wells, 2007) approximately 30kya. This marker, which is found at 25% in north Afghanistan and 4.8% in the south, has also been previously reported at high frequencies (48%) in the Kallar community of South India (Wells *et al*., 2001) as well as in the Druze (35%) population from Israel.( Shen *et al*., 2004) Time estimates generated based on seven Y-STR loci within L-M20 lineages for north (14.6±7.3kya) and south (17.8±8.4kya) Afghanistan populations are intermediate to those of Pakistan (26.3±5.3kya) and India (7.5±1.7kya) (Lacau *et al*., 2012). Furthermore, Pakistan displayed higher haplotype variance (0.548) than India (0.118), suggesting that L-M20 most likely originated in what is today Pakistan rather than in India (Lacau *et al*., 2012) and findings of present investigation confirms this argument and also support the dispersal of this haplogroup L from its place of origin to Southern Afghanistan.

Likewise, the highest frequencies of Y-chromosome Haplogroup H-M69 are in India and Pakistan especially in 4.1% Burusho, 20.5% Kalash, 4.2% Pathan and 6.3% in other Pakistani populations (Thanseem *et al.,* 2006; Qamar *et al*., 2002). It is also present in high frequency of 25 – 35% in tribal population and in low frequency of 10% in Western India and Pakistan (Sahoo *et al*., 2006; Cordaux *et al*., 2004).The low frequency of Ychromosome Hg R1b, E1b1b and J2 of predominantly Greek lineage in these populations and in the Pathan population of Afghanistan was observed, which strengthen the idea of Firasat *et al*., (2007) about the limited Greek contribution to the Y chromosome structure of Pathans. The E1b1b1-M35 lineages in some Pakistani Pathan were previously traced to a Greek origin brought by Alexander" s invasions (Firasat *et al*., 2007).

For mtDNA analysis, it has been noticed that profound genetic variations were accounted within populations (97.97%) rather than between the populations (3.63%). while variance distribution was greater within populations (1.3%) than among the populations (0.04%). This relatively high heterogeneity among these Pathan populations further modeling the mtDNA landscape that might be the result of unequal genetic influence obtained during different settlement events from geographically adjacent neighbors. On account of that, some haplotypes are shared across the tribal boundaries which may be due to maledriven intermarriages.

The mtDNA Macro haplogroups recovered from Pathans of Mardan and Charsada are R, M, N and L. The first three haplogroups are thought to be originated around 60000-75000 years ago in South Asia (Kivisild, 2003). Therefore, the presence of these haplogroups suggests a South Asian maternal origin of these populations. The macro-haplogroup L (Sub-Saharan Africa) was detected in just 0.5% i.e. only on one subject, while East African haplogroups (L4, L5, L6 and L7) are also not detected in Pathans as yet; while all the Arabian Peninsula countries are more likely to have a comparatively high frequency, such as Yemen (38%), Oman (16%) and Saudi Arabia (10%) (Abu-Amero *et al*., 2007). The most common South Asian haplogroups in Pathans comprises of R (67.9%) and M (22.3%). In the present study, the mitochondrial haplogroups of South Asian origin have highest proportion of 37.6%, comprising M3 (7.5%), R2(5.3%), U7, U2e (4.3%), M4, U6c (3.2%), U2c, M65a, M6a1b, U5a1b, M49, M5, M2a1a, M6a1b, R0a, U2e1h (2.2%) and U5a2, U2e3, U4a3, U2e1f, U2b2, U2a1a, U2a, M52a, M5a1, M33, M38d, M39a, M2b, M18b,

M71(1.1%). The second major haplogroups identified were of Southwest Asian or Middle Eastern origin (36.5%) including H2a2a1g (6.4%), J1b1a1, H2 (5.2%), H2a2a1c, H5 (4.3), H10, R6 (3.2), H15a, H1e, H2a2a1d, R8a1a3, J1b, J1b5b (2.2%). The third most prevalent mitochondrial haplogroups are of Central Asian origin (8.6%) followed by West Eurasian (6.5%), east Asian and Indo-Europeans (3.2%) and the least frequent haplogroup was found L (1.1%) of East African origin. The high frequency of macro haplogroup M had been recorded in Asia, particularly in India, Bangladesh, Nepal and Tibet reaching up to 60 to 80% (Rajkumar *et al*., 2005). Mitochondrial haplogroup M3 being the sub-clad of HgM, has been found in South Asia, with highest frequency in West India and Pakistan (Metspalu *et al*., 2004). It has been reported in Pakistani Pathan population with frequency of 7.8% (Rakha *et al*., 2011) and this study reported the 7.5% occurrence of this haplogroup.This concurrence of mtDNA markers could evident significant female enigmatic contribution towards the genetic structuring and socio-cultural expansion of multiethnic Pathan population (Rakha *et al*., 2011). Moreover this uniparental genetic landscape of mtDNA HVI data proposed that the Pathans revealed a profound blend of genetic components from South Asia, the Caucasus, Europe, Arabian Peninsula and Central Asia (Quintana-Murci *et al*., 2004). It is obvious that the majority of Pathan autochthonous M lineage huddled into M5 and M3 sub-haplogroups which have characteristic South Asian origin.

Phylogeographic analysis has revealed that M5a1 haplogroup is most abundant in all Roma Gypsies population, including Hungarian, Bulgarian, Iberian and Balkan Roma (ranging from 6 to 29%), constituted a founder population that distributed across West

Asia and whose origin might be traced back to the Indian Subcontinent, particularly Pakistan (Mendizabal *et al*., 2011). In these five populations it was found in just 1.1%. Another mitochondrial DNA lineage, the West Eurasian macrohaplogroup N has been proved to be the ancestor of numerous haplogroups prevailing in Europe, Middle East, Asia and the Americas (Kivisild *et al*., 1999; Nasidze *et al*., 2008). It is present though less common in Pathans, with the overall frequency of 8.6% in their genetic pool, which is in agreement with previous studies (Quintana-Murci *et al*., 2004; Rakha *et al*., 2011). The haplogroup R (sub-clad of N) cuddled the majority of the West Asian and European haplogroups H, HV, J, T and U, which have varying frequencies in fivePathan populations (Malmstrom *et al*., 2015). The incidence of overwhelming immensity of R lineage (63.4%) in Pathan has a clear Europe and West Asia attribution due to Paleolithic and Neolithic expansions of Caucasian that reached in South Asia via Iranian plateau and might be Arabian Sea nautical routes, could be the possible argument for this genetic influx (Stoljarova *et al*., 2016; Haber *et al*., 2016). Furthermore, haplogroup R0 comprising HV0, HV1 and HV2 was scarcely found in these fivePathanpopulations (1.15%) as compared to other Pakistani population, having highest frequency in Balochi (10.3%), Pathans (9.1%) and absent in Burusho and Hazara (0%) (Kivisild *et al*., 2002; AlvarezIglesias *et al*., 2009). However, neighboring countries exhibited low frequency such as Uzbek (2.4%), Turkmen (2.4%) and in Turkish its frequency dropped to 0% (Rakha *et al*., 2011; Siddiqi *et al*., 2015). The haplogroup H was 9.8% prevalent in overall five Pathan populations but different from West Asia (45%) and Near East (25%) (Palanichamy *et al*., 2015). Whereas neighboring population exhibit relatively high frequency of this haplogroup, such as Iranian (14.3%),

Uzbek (21.4%), Turkmen (22%) and Tajik (29.5%) (Whale,2012; Derenko *et al*., 2013). Similarly other Pakistani populations exhibited very high frequency of haplogroup H i.e. Sindhi (28%), Brahui (26.3%), Balochi (20.5%) but moderate in Burusho (12.3%) and Hazara (13%) population (Quintana-Murci *et al*., 2004; Bhatti *et al*., 2016). Notably, under these assumptions it is hard to discern sequential gene flow or expansions at the population level because the most recent migration i.e. British colonization in the 18th century could hold both early and derivative lineages, which contributed significantly in the genetic pool of the Pathan (Jons, 2016). Besides aforementioned haplogroups, analysis of specific haplogroups J revealed some interesting geological and chronological distribution in Pathan. It was found with the frequency of 9.6% reflected a deep Jewish and European ancestry conglomeration. It is noteworthy that expansion of the Jewish community particularly Ashkenazi, across the Roman Empire and Iranian Plateau, was going to be around 1000 BC (Mosk, 2013). However, on account of expulsion in the Western Europe during the 15th century, they subsequently dispersed into Eastern Europe and moved up to Afghanistan, North West Pakistan and India (Behar *et al*., 2004; Shlush *et al*., 2008; Costa *et al*., 2013). Yet, Pakistani and Indian Jewish communities are an early branch of the Jewish Diaspora with their several exclusive socio-cultural features, intricate history and minor Middle East specific ancestry components (Oddie, 1991; Ostrer, 2001; Chaubey *et al*., 2016). Another distinguished mitochondrial super-haplogroup U that was most abundant in Pathan (21.8%), which includes haplogroups U2, U5 and U7.

The haplogroup U was widely distributed in the Europe and Near East with the coalescence age of about 51,000–67,000 years before present (Roostalu *et al*., 2007).

However, this widespread distribution, when incorporated with the presence of a profound associated haplogroup U2 found in native South Asians, suggests that haplogroup U2 is very much entrenched, divergent and was presumably a former lineage of the super-haplogroup U in South Asia (Barcaccia *et al*., 2015). This statement was further strengthened by the incidence of U2 haplogroup in excavated human remains in the Southern Russia at around 30 000 years before present i.e. during Pleistocene expansion. Due to frequent distribution of U2 in South Asian and European huntergatherer lineages, it has practically predicted some association among „„Paleolithic Hunters" " and present day native inhabitants (Krause *et al*., 2010). Another aforementioned mitochondrial lineage, U7 has originated in the „„Black Sea" " zone between Southeastern Europe and Western Asia dated back to 51 000–67 000 years before present. It is markedly found among India (20%) and Near East (10%) (Roostalu *et al.,* 2007) and found 4.3% in these Pathan, 23% in North Indians, 10% in South Indians, 7.2% in Iran while absent in the Western Asian or Anatolian and Eastern European or Balkans population (Palanichamy *et al*., 2015; Bahmanimehr *et al*., 2015; Chaitanya *et al*., 2015). This implied some degree of complex overlapping of haplogroups, across the entire Persia, Pakistan (KPK) and India during various migratory proceedings that was an outcome of succeeding historical and geographic gene flow (Bahmanimehr *et al*., 2015). Despite primary sub-haplogroups of U, one European-specific haplogroup U5 exhibited relative frequency of 3.3% in Pathan. Earlier investigations of ancient mtDNA proposed that U5 haplogroup was most frequent in Mesolithic and Neolithic Europeans. For example, a high frequency (65%) of U5

haplogroups has been assessed in the European hunter-gatherers individuals (Malmstrom *et al*., 2009; Sanchez-Quinto *et al*., 2012; Cassidy *et al*., 2016).

The frequency of Y-haplogroups and mtDNA haplogroups on the basis of their origin in these Pathans is almost similar with a little difference. The plausible explanation for this difference may be the limited sample size typed for mtDNA analysis and a different picture could be expected by increasing the sample size for populations.

The frequency of these mtDNA macro haplogroups found in the populations we genotyped in present study is in close agreement with the one already reported in Pakistani Pathans with a little variation in frequency of sub-clades. Mitochondrial haplogroups HV9, M65, H1k, U2a1, M71, M52 and U5a1b were confined to the populations of Mardan (Mohmand: MM, Muhammadzai and Kakakhel Mian) only while haplogroups U7, M18 and P6 were restricted to Charsada populations (Yousafzai and Mohmand: MD) only and absent in Mardan. The high frequencies of local lineages in these populations confirm their descent as a consequence of Paleolithic population expansion (McElreavey *et al*., 2005).

It is indicative that the pattern of Y-haplogroup distribution in Afghanistan and already studied general Pathan population from Pakistan investigated by Sangupta *et al*., 2006, show great similarity. In the light of our results and their homology with the previously documented genetic information, it can be inferred that these populations have been descended from Durrani Pathans in South and Southwest Afghanistan which is in agreement with the oral historical tradition that the potential paternal ancestors of these

Pathan populations were Durrani which were descendants of haphthalites rooted from Qais (Allah *et al*., 2013). On mtDNA-$F_{ST}$ based PCoA scatter plot, Mohmand from Charsada (MM) showed great genetic affinity with Turkey, Israel and Iran positioned somewhere at the middle of these three, depicting their admixed maternal lineage from Turkey, Israel and Iran rather than Afghanistan. Similarly Yousafzai and Mohmand (MD) from Mardan show relatedness with Afridi Pathans from West India. The logical explanation for different clustering on mtDNA variation may be the different origin of matrilines and patrilines of these populations. But the difference between paternal and maternal origin is not confirmed by although very little positive correlation between distance matrix of mtDNA and Y-STR. In the light of great genetic similarity revealed among the populations studied, results validate the origin of the three of the four Pathan tribes from the single patriline as documented in the legend, i.e. from "Sarban confederacy" While lack of Arabian Y clusters and mtDNA haplogroups in Kakakhel Mian population refute the Arabian descent of Kakakhel Mian. Another logical explanation for different clustering pattern may be the more frequent male movement among the two areas as compared to the females which contribute to the more gene flow in paternal gene pool reducing the differentiation among the males of these populations while strict practice of endogamy within the populations contribute to retain their genetic differentiation on matrilines.

## CONCLUSION

On the basis of our findings compared with already available scientific information on

Pathans we conclude that vast majority of paternal gene pool of the Pathans from Mardan and Charsada populations exhibit genetic uniformity with other Pathans from Pakistan and Afghanistan, displaying a mosaic gene pool with haplogroup sharing of the Middle Eastern, Caucasian, South Asian and Central Asian gene pools. It can be inferred from our results that these populations are donors of the Pathan population of areas of Afghanistan from where these populations migrated to their present settlements. Furthermore the Pathans have maintained the Genetic diversity irrespective of the geographical location and sub-population or caste. The possible reason for this pattern is the common practice of endogamy restricting their intermixing with other ethnic groups. Moreover our results are congruent with the above stated hypothesis about the legendry Afghan ancestor, where these populations exhibit considerable homogeneity contributed by being a descendant of a single Y chromosome i.e. Sarban. The study presented in this dissertation reports the first hand information with respect to Pathans of KP province which can be extended to other areas of the region with maximum analytical material and tools for establishing intra and inter population affinities of the tribes and clans. This information can be used as a useful database for establishing a forensic baseline for the Pathan population of the region. For complete elaboration of the phylogenetic position of Pathan lineage, complete Mitogenome sequence analysis and Y chromosomal SNP analysis is required.

## REFERENCES

Alakoc, Y. D, O. Gokcumen, A. Tug, T. Gultekin, E. Gulec, T.G. Schurr. 2010. Y-

chromosome and autosomal STR diversity in four proximate settlements in Central Anatolia. *Forensic Sci Int Genet* 4: e135–e137

Ali, I. 1994. Settlement history of Charsadda District. Ancient Pakistan, 9:175.

Ali, S and S. E. Hasnain. 2003. Genomics of the human Y-chromosome 1. Association with male infertility, Gene,321: 25-37.

Allah, H. N. and S. M. Imamuddin. 1960. Tarikh-i-Khan Jehani wa Makhzan-i-Afghani. Dacca, Asiatic society of Pakistan. Publ. No. 4. Vol. 2

Allah, H.N. and B. Dorn. 2013. History of the Afghans: Translated from the Persian of Neamet Ullah. Cambridge University Press.

Allen, M., A. S. Engstrom, S. Meyers, O. Handt, T. Saldeen, *et al*. 1998. Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. *Journal of Forensic Sciences*, 43: 453-464.

Alonso, A., A. Salas, C. Albarran, E. Arroyo, A. Castro, *et al*. 2002. Results of the 19992000 collaborative exercise and proficiency testing program on mitochondrial DNA of the GEP-ISFG: an interlaboratory study of the observed variability in the heteroplasmy level of hair from the same donor. *Forensic Science International*, 125: 1-7.

Alvarez-Iglesias, V., A. Mosquera-Miguel, M. Cerezo, B. Quintans, M. T. Zarrabeitia, I. Cusco, M. V. Lareu, O. Garc ıa, L. Perez-Jurado, A.Carracedo, 2009. New population and phylogenetic features of the internal variation within mitochondrial DNA macro-haplogroup R0. PLoS One. 4:e5112.

Ambrose, S. 1998. Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. *Journal of Human Evolution*, 34(6): 623–651.

Andrews, R. M., I. Kubacka, P. F. Chinnery, R. N. Lightowlers, D. M. Turnbull, N. Howell. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet.*, 23:147

Asari, M., K. Umetsu, N.Adachi, J.I.Azumi, K. Shimizu and H.Shiono, 2007.Utility of haplogroup determination for forensic mtDNA analysis in the Japanese population. *Legal Medicine*, *9*(5): 237-240.

Awde, N. and A. Sarwan. 2002. Pashto: Dictionary & Phrasebook. New York: Hippocrene Books Inc

Awde, N. and A. Sarwan. 2003. Pashto: Pashto-English, English-Pashto.Dictionary & Phrasebook. Hippocrene Books.

Babur. 1987. Babur Nama (Memories of Babur). Vol. II, Eng. Trans. A. S. Beveridge, Lahore.

Balamurugan, K., G. Suhasini, M. Vijaya, S. Kanthimathi, Mullins, M. Tracey, G. Duncan. 2010. Y chromosome STR allelic and haplotype diversity in five ethnic Tamil populations from Tamil Nadu, India. *Int J Legal Med.,* 12: 265–269

Ballantyne, K. N., M. Goedbloed, R. Fang, O. Schaap, O. Lao, A. Wollstein, Y. Choi, K. van Duijn, M. Vermeulen, S. Brauer, R. Decorte, M. Poetsch, N. Von Wurmb-Schwark, P. de Knijff, D. Labuda, H. Vézina, H. Knoblauch, R. Lessig, L. Roewer, R. Ploski, T. Dobosz, L. Henke, J. Henke, M. R. Furtado, M. Kayser. 2010.

Mutability of Y chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *Am. J. Hum. Genet*., 87: 341–353.

Ballantyne, K. N., V. Keerl , A. Wollstein , Y. Choi , S. B. Zuniga , A. Ralf , M. Vermeulen , P. de Knijff , M. Kayser . 2012. A new future of forensic Ychromosome analysis: rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. *Forensic Sci Int Genet.,* 6:208-218.

Balter, M. 2010. Of Two Minds About Toba's Impact. *Science,* 327 (5970): 1187–1188.

Banuazizi, A. and M. Weiner (eds.). 1994. The Politics of Social Transformation in Afghanistan, Iran, and Pakistan. Syracuse University Press. pp.488

Barcaccia, G., G. Galla, A. Achilli,A. Olivieri, A.Torroni, 2015. Uncovering the sources of DNA found on the Turin Shroud. *Scientific Reports*, 5:14484.

Battaglia V., S. Fornarino, N. Al-Zahery, A. Olivieri, M. Pala, N. M. Myres, R. J. King, S. Rootsi, D. Marjanovik, D. Primorac, R. Hadziselimovic. 2009. Y-chromosomal Evidence of the Cultural Diffusion of Agriculture in Southeast Europe. European Journal of Human Genetics, 17(6): 820-830

Behar, D.M., M.F.Hammer, D.Garrigan, R.Villems, B.Bonne-Tamir, M.Richards, D. Gurwitz, D.Rosengarten, M.Kaplan, S.Della Pergola, and L.Quintana-Murci, 2004.MtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. *European Journal of Human Genetics*, *12*(5): 355-364.

Bhatti, S., M. Aslamkhan, M. Attimonelli, S. Abbas, and H. H. Aydin, 2016. Mitochondrial DNA variation in the Sindh population of Pakistan. *Australian Journal of Forensic Sciences*, pp.1-16.

Blood, P. R., 2001. Afghanistan: A Country Study:"The Reign of King Habibullah, 19011919", The Reign of King Amanullah, 1919-29","Reform, Popular Reaction, and

Forced Abdication","Tajik Rule, January-October 1929" and "Muhammad Nadir Shah, 1929-33". US Library of Congress.

Bower, B., 2011. Story one: Stone tools hint at earlier human exit from Africa: Controversial find could push expansion back 65 millennia. *Science News*, 179(5): 5-6.

Bowler, J. M., H. Johnston, J. M. Olley, J. R. Prescott, R. G. Roberts, W. Shawcross, N. A. Spooner. 2003. New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature*, 421 (6925): 837–40.

Brown, W. M., M. J. George, A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. Proc. *Natl. Acad. Sci. USA*. 76(4): 1967-71.

Budowle, B., 1995. The effects of inbreeding on DNA profile frequency estimates using PCR-based loci. In *Human Identification: The Use of DNA Markers*. Springer Netherlands, 4: 21-25.

Burton, M. L., C. C. Moore, A. K. Romney, D. F. Aberle, J. A. Barcelo, M. M. Dow, J. I. Guyer, D. B. Kronenfeld, J. E. Levy, J. Linnekin, 1996. Regions based on social structure. *Curr. Anthropol.*, 37(1): 87-123.

Butler, J. M. 2003. Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis.*Forensic SciRev.,* 15(2): 91-111.

Butler, J. M. 2012. Advanced Topics in Forensic DNA Typing: Methodology.Academic Press, p.704.

Butler, J. M., A. E. Decker, M. C. Kline, and P. M. Vallone. 2005. Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation. *Journal of Forensic Sciences*, 50(4):853–859.

Butler, J.M., E. Buel, F. Crivellente, B. R. McCord. 2009. Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis,* 25(10-11):1397- 412.

Callaway, E. 2015.Teeth from China reveal early human trek out of Africa. Stunning" find shows that *Homo sapiens* reached Asia around 100,000 years ago.*Nature*,News, 18566.

Cann, R. L., M. Stoneking, A. C. Wilson. 1987. Mitochondrial DNA and human evolution. *Nature*, 325: 31-36.

Caroe, O., 1958.The Pathans: 550 BC – AD. 1957. Oxford University press Karachi.

Caroe, O., 1984. The Pathans: 500 BC-AD. 1957. (Oxford in Asia Historical Reprints).

Cassidy, L.M.,R. Martiniano, E.M.Murphy, M.D.Teasdale, J.Mallory, B.Hartwell, and

D.G.Bradley, 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proceedings of the National Academy of Sciences*, *113*(2): 368-373.

Cavelier, L., A. Johannisson, U. Gyllensten. 2000. Analysis of mtDNA copy number and composition of single mitochondrial particles using flow cytometry and PCR. *Experimental cell research*, 259(1): 79- 85.

Chaitanya, L., A. Ralf, M. Oven, T. Kupiec, J. Chang, R. Lagacé, and M. Kayser, 2015. Simultaneous whole mitochondrial genome sequencing with short overlapping amplicons suitable for degraded DNA using the ion torrent personal genome machine. *Human mutation*, *36*(12): 1236-1247.

Chandrasekar, A., S. Kumar, J. Sreenath, B. N. Sarkar, B. P. Urade, S. Mallick, S. S. Bandopadhyay, P. Barua, S. S. Barik, D. Basu, and U. Kiran, 2009. Updating phylogeny of mitochondrial DNA macrohaplogroup m in India: dispersal of modern human in South Asian corridor. *PloS one*, *4*(10): e7447.

Chaubey, G., M. Singh, N. Rai, M. Kariappa, K. Singh, A. Singh, D. P. Singh, R. Tamang, D. S. Rani, A. G. Reddy and V.K.Singh, 2016. Genetic affinities of the Jewish populations of India. *Scientific reports*, *6*:19166.

Chen, Y. S., A. Torroni, L. Excoffier, A. S. Santachiara-Benerecetti, D. C. Wallace. 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *American journal of human genetics*, 57(1): 133-149.

Chen, Y. S., A. Torroni, L. Excoffier, A. S. Santachiara-Benerecetti, and D. C. Wallace, 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *American journal of human genetics*, *57*(1): 133.

CIA. July 2012. Country Comparison: Population. Retrieved 2 September 2012

Comas, D., F. Calafell, E. Mateu, A. Pe" rez-Lezaun, J. Bertranpetit. 1996. Geographic Variation in Human Mitochondrial DNA Control Region Sequence: The Population History of Turkey and its Relationship to the European Populations. *Mol. Biol.,* 13(8):1067-1077.

Comas, D., S. Plaza, R. S. Wells, N. Yuldaseva, O. Lao, F. Calafell and J. Bertranpetit, 2004. Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. *European Journal of Human Genetics*, *12*(6): 495-504.

Cordaux, R., R. Aunger, G. Bentley, I. Nasidze, S. M. Sirajuddin, M. Stoneking. 2004. Independent Origins of Indian Caste and Tribal Paternal Lineages. *Current Biology*, 14 (3): 231–235.

Costa, M. D., J. B. Pereira, M. Pala, V. Fernandes, A. Olivieri, A. Achilli, U. A. Perego, S. Rychkov, O. Naumova, J. Hatina, and S. R. Woodward, 2013. A substantial prehistoric European ancestry amongst Ashkenazi maternal lineages. *Nature Communications*, *4*.

Cruciani, F., P. Santolamazza, P. Shen, V. Macaulay, P., Moral, A. Olckers, D. Modiano, S. Holmes, G. Destro-Bisol, V. Coia, D.C. Wallace. 2002. A Back Migration from Asia to Sub-Saharan Africa Is Supported by High-Resolution Analysis of Human Y-Chromosome Haplotypes. *American Journal of Human*

*Genetics,* 70(5): 1197–1214.

Cummins, J. 1998. Mitochondrial DNA in mammalian reproduction. *Reviews of Reproduction*, 3: 172-182

Cummins, J. M. 2000, Fertilization and elimination of the paternal mitochondrial genome, *HumanReproduction*, 2: 92-101

Cunningham, A., 1963. The Ancient Geography of India: The Buddhist period, including the campaigns of Alexander, and the travels of Hwen-Thsang. I. (Vol. 1). Genesis Publishing Pvt Ltd.

Dani, A. H. 1963. Pushkalavati: The Lotus City, Peshawar University Archeological Guide Series No.1, Peshawar.

Dennell, R. and M. D. Petraglia. 2012. The dispersal of Homo sapiens across southern Asia: how early, how often, how complex?. *Quaternary Science Reviews*, 47: 15 -22.

Derenko, M., B. Malyarchuk, G. Denisova, M. Perkova, A. Litvinov, T. Grzybowski, I. Dambueva, K. Skonieczna, U. Rogalla, I. Tsybovsky and I. Zakharov, 2014. Western Eurasian ancestry in modern Siberians based on mitogenomic data. *BMC evolutionary biology*, *14*(1): 217.

Derenko, M., B. Malyarchuk, A. Bahmanimehr, G. Denisova, M. Perkova, S. Farjadian, L. Yepiskoposyan. 2013. Complete Mitochondrial DNA Diversity in Iranians. *PLoS one*, 8(11): e80673.

Di Cristofaro, J., E. Pennarun, S. Mazières, N. M. Myres, A. A. Lin, S. A. Temori, M. Metspalu, E. Metspalu, M. Witzel, R. J. King, P. A. Underhill, R. Villems, J.

Chiaroni. 2013. Afghan Hindu Kush: Where Eurasian Sub-Continent Gene Flows Converge. *PLoS one*, 8(10): e76748.

Di Cristofaro, J., S. Buhler, S. A. Temori, and J. Chiaroni, 2012.Genetic data of 15 STR loci in five populations from Afghanistan. *Forensic Science International: Genetics*, *6*(1): e44-e45.

Disotell, T.R., 1999. Human evolution: the southern route to Asia. *Current Biology*, *9*(24): R925-R928.

District Census Report of Charsadda. 1998. Census Publication No. 68. Islamabad: Population Census Organization Statistics Division, Government of Pakistan.

Dupree, L., J. L. Angel, R. H. Brill, E. R. Caley, R. S. Davis, C. C. Kolb, A. Marshack, D. Perkins and A. Solem, 1972. Prehistoric research in Afghanistan (1959-1966). *Transactions of the American Philosophical Society*, *62*(4): 1-84.

Dupree, N. H. 1977. An Historical Guide to Afghanistan. Kabul: Afghan Air Authority, Afghan Tourist Organization publication.

Dupuy, B. M., M. Stenersen, T. Egeland and B. Olaisen. 2004. Y-chromosomal microsatellite mutation rates: Differences in mutation rate between and within loci, *Human Mutation.* 23:117–124.

Ellegren, H. 2000. Heterogeneous mutation processes in human microsatellite DNA sequences. *Nat. Genet.*, 24:400-402.

Elphinstone, M. 1842. An Account of the Kingdom of Caubul, and its Dependencies in Persia, Tartary and India. Vol. II. London: RichardRentley, New Burlington Street.

Emadi H., 2005. Culture and Customs of Afghanistan. Santa Barbara, CA: Greenwood. p .284.

Excoffier, L. and H. E. L. Lischer. 2010. Arlequin suite ver. 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Eco. Res*., 10: 564-567.

Fernandes, A. T., R. Gonçalves, S. Gomes, D. Filon, A. Nebel, M. Faerman, A. Brehm. 2011. Y-chromosomal STRs in two populations from Israel and the Palestinian Authority Area: Christian and Muslim Arabs. *Forensic Science International: Genetics*, 5: 561–562.

Firasat, S., S. Khaliq, A. Mohyuddin, M. Papaioannou, C. T. Smith, P. A. Underhill, Q. Ayub. 2007. Y-chromosomal evidence for a limited Greek contribution to the Pathan population of Pakistan.*European Journal of Human Genetics*, 15(1): 121–126.

Foreman, L. A. and J. A. Lambert. 2000. Genetic differentiation within and between four UK ethnic groups. *Forensic Science International,* 114(1): 7–20.

Gankovsky, Yu. V. 1982. A History of Afghanistan. Moscow: Progress Publishers, p. 382

Garrick, H. W. B. 1882. Report on tour through Behar, Central India, Peshawar and Yousafzai, Archeological Survey of India Report Vol. XIX, 94-110.

Gazetteer of Peshawar District (1897-98). 1989. Sang-e-Meel Publications, Lahore.

Ghosh, T. D., S. Kalpana, M, Mukerjee, A. K. Mukherjee, S. Sharma, V. R. Nath, M. K. Rathod, G. N. Thakar. 2011. The Genetic diversity of 17 Y-short tandem repeats in Indian population, *Forensic Sci. Int., Genetics* 5(4): 363–367.

Giles, R. E., H. Blanc, H. M. Cann, D. C. Wallace. 1980. Maternal inheritance of human mitochondrial DNA. Proc. *Natl. Acad. Sci. USA*., 77: 6715-19.

Gill P., L. Gusmão, H. Haned, W. R. Mayr, N. Morling, W. Parson, L. Prieto, M. Prinz, H. Schneider, P.M. Schneider, B.S. Weir. 2012. DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods. *Forensic Sci. Int., Genetics* 6(6): 679-688.

Gopaldas, M. 1874. Tareekh-i-Peshawar, Urdu, Lahore. P: 273-74.

Gopaldas, M. 1874. Tehreek-e-Peshawar. p. 319.

Gusmão, L., P. Sánchez-Diz, F. Calafell, P. Martin, C. A. Alonso, F. Álvarez Fernández, C. Alves, L. Borjas Fajardo, W. R. Bozzo, M. L. Bravo, J. J. Builes. 2005. Mutation rates at Y chromosome specific microsatellites. *Human Mutation*, 26(6): 520–528.

Gusmão, L., M. Brion, A. González-Neira, M. Lareu, A. Carracedo. 1999. Y-chromosome specific polymorphisms in forensic analysis. *Legal Medicine,* 1(2): 55-60. Haber, M., M. Mezzavilla, Y. Xue and C. Tyler-Smith, 2016. Ancient DNA and the rewriting of human history: be sparing with Occam" s razor. *Genome biology*, *17*(1):

1.

Hallast P., C. Batini, D. Zadik, P. M. Delser, J. H. Wetton, E. Arroyo-Pardo, G. L. Cavalleri, P. De Knijff, G. D. Bisol, B. M. Dupuy, H. A. Eriksen. 2015. The YChromosome Tree Bursts into Leaf: 13,000 High-Confidence SNPs Covering the Majority of Known Clades. *Mol. Biol. Evol.*, 32(3): 661–673.

Hammer, M. F. 1995. A recent common ancestry for human Y chromosomes. Nature, 378(6555): 376–378.

Hammer, M. F. and S. L. Zegura. 2002. The human Y chromosome haplogroup tree: Nomenclature and Phylogeography of its major divisions. *Annual Review of Anthropology*, 31: 303–321.

Hammer, M.F. 1994. A recent insertion of an Alu element on the Y-chromosome is a useful marker for human population studies. *Molecular Biology Evolution,* 11(5): 749-761.

Hanson, E. K. and J. Ballantyne. 2006. Comprehensive annotated STR physical map of the human Y-chromosome: Forensic implications. *Legal Medicine,* 8(2): 110-120.

Haroon, S. 2007. Frontier of Faith: Islam, in the Indo-Afghan Borderland. Columbia University press New York. P.40.

Harpending, H. and G. Cochran. 2009. The 10,000 Year Explosion: how civilization accelerated human evolution. Basic Books. p. 214.

Harvey, S. L., W. P. Lehmann and J. Slocum. 2003-2004. Old Iranian Online: Introduction and 10 lessons. University of Texas, Austin

Hastings, E. G. G. 1878. Report of the Regular Settlement of the Peshawar District of the Punjab. Lahore: Central Jail Press, 103-108.

Hayat, S., T. Akhtar, M. H. Siddiqi, A. Rakha, N. Haider, M. Tayyab, G. Abbas, A. Ali, , S. Y. A. Bokhari,M. A. Tariq and F. M. Khan, 2015. Mitochondrial DNA control region sequences study in Saraiki population from Pakistan. *Legal Medicine*, *17*(2): 140-144.

Hedman, M., A. Brandstätter, V. Pimenoff, P. Sistonen, J. U. Palo, W. Parson and A. Sajantila, 2007. Finnish mitochondrial DNA HVS-I and HVS-II population data. *Forensic Science International*, *172*(2): 171-178.

Helena, M. A. and B. J. Morris. 2007. The Human Pseudoautosomal Region (PAR): Origin, Function and Future. *Current Genomics,* 8(2): 129-136.

Holdich,T. H. 1899. Swatis and Afridis.The Journal of the Anthropological Institute of Great Britain and Ireland, 29 (1/2): 2-9.

Howell, N., J. L. Elson, C. Howell, D. M. Turnbull. 2007. Relative rates of evolution in the coding and control regions of African mtDNAs. *Molecular biology and evolution,* 24: 2213-2221.

Hubert, R., J. L. Weber, K. Schmitt, L. Zhang and N. Arnheim. 1992. A new source of polymorphic DNA markers for sperm typing: Analysis of microsatellite repeats in single cells. *Am J Hum Genet*., 51 (5): 985–91.

Hunley, K. L., M. E. Healy, J. C. Long. 2009. The Global Pattern of Gene Identity Variation

Reveals a History of Long-Range Migrations, Bottlenecks, and Local Mate Exchange: Implications for Biological Race. *Am J phys Anthropol*., 139:35– 46

Hurles, M. E. and M. A. Jobling. 2001. Haploid chromosomes in molecular ecology: lessons from the human Y. *Molecular Ecology,* 10: 1599-1613.

Hussain, J. 1992. Pottery Classification System- A Proposed Model for Shaikhan Dheri pottery. *Pakistan archeology*, 25. Karachi

Ilyas, M., M. S. Shahzad, M. Israr, S. S. Jafri, M. Shafeeq, M. S. Zar, A. Ali, Z. Rahman, T. Husnain. 2012. Y-Chromosomal STR Haplotype Profiling in Yousafzai" s living in Swat Valley Pakistan. Presentations from the 22nd Congress of the IALM *Istanbul, Turkey*

Imad, H., A. Muhanned, J. Aamera, Y. Cheah. 2014. Analysis of eleven Ychromosomal STR markers in middle and south of Iraq. *Afr. J. Biotechnol*. 13(38): 3860-3871.

Ingman, M., H. Kaessmann, S. Pääbo, U. Gyllesten. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature*, 408(6813): 708-713.

Ingman, M., U. Gyllensten. 2001. Analysis of the complete human mtDNA genome: methodology and inferences for human evolution. *The Journal of heredity*, 92(6): 454-461.

Irwin, J., J. Saunier, K. Strouss, C. Paintner, T. Diegoli, K. Sturk, L. Kovatsi, A. Brandstatter, M. A. Cariolou, W. Parson, T. J. Parsons. 2008. Mitochondrial control region sequences from northern Greece and Greek Cypriots.*Int. J. Legal Med*., 122 (1):87-89.

Jeffreys, A. J., J. Murray and R. Neumann. 1998. High-resolution mapping of crossovers

in human sperm defines a minisatellite-associated recombination hotspot. *Mol. Cell,* 2(2): 267–273.

Jobling, M. A. and C. Tyler-Smith. 2001. Y- chromosomal SNP haplotype diversity in forensic analysis. *Forensic Science International,* 118(2): 158-162.

Jöns, H., 2016. The University of Cambridge, academic expertise and the British empire, 1885–1962. *Environment and Planning A*, *48*(1): 94-114.

Kakakhel, S. B. S. Z. 1981. Pashtoon Taareekh Kay Aienay Main (Pashtoons in the light of history), Abdur Rasheed Press Gujrat.

Karafet, T., F. L. Mendez, M. B. Meilerman, P. A. Underhill, S. L. Zegura and M. F. Hammer. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research,* 18 (5): 830–8.

Karlsson, A. O., T. Wallerström, A. Götherström and G. Holmlund. 2006. Y chromosome diversity in Sweden – A long-time perspective. *Eur. J. Hum.Genet.,* 14 (8): 963–970.

Katzir, I., 2001. The Israeli source of the Pathan tribes. *The Scribe Magazine*, *74*: 53-54.

Kayser, M. and A. Sajantila. 2003. Mutations at Y-STR loci: implications for paternity testing and forensic analysis. *Forensic Science International,* 118(2): 116-121.

Kayser, M., S. Brauer, G. Weiss, W. Schiefenhövel, P. Underhill, P. Shen, P. Oefner, M. Tommaseo-Ponzetta, M. Stoneking. 2003. Reduced Y-chromosome, but not mtDNA, diversity in human populations from West New Guinea. *Am. J. Hum.Genet.*, 72(2):281–302.

Kayser, M., S. Brauer, H. Schadlich, M. Prinz, M. Batzer, P. Zimmerman, B. Boatin and M. Stoneking. 2003. Y chromosome STR haplotypes and the genetic structure of U.S. populations of African, European, and Hispanic ancestry. *Genome Res*., 13(4): 624–634.

Kayser, M., S. Brauer, G. Weiss, W. Schiefenhövel, P. A. Underhill and M. Stoneking. 2001. Independent histories of human Y chromosomes from Melanesia and Australia. *Am. J. Hum. Genet.,* 68(1): 173–190.

Kivisild,T., M. J. Bamshad, K. Kaldma, M. Metspalu, E. Metspalu, M. Reidla, S. Laos, J. Parik, W. S. Watkins, M. E. Dixon and S. S. Papiha, 1999. Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Current Biology*, *9*(22): 1331-1334.

Kivisild, T., S. Rootsi, M. Metspalu, S. Mastana, K. Kaldma, J. Parik, E. Metspalu, M. Adojaan, H. V. Tolk, V. Stepanov, M. Gölge, E. Usanga, S. S. Papiha, C. Cinnioğlu, R. King, L. Cavalli-Sforza, P. A. Underhill, R. Villems, 2003. The Genetic Heritage of the Earliest Settlers Persists Both in Indian Tribal and Caste Populations. *Am. J. Hum.Genet.,* 72(2): 313–332.

Kong, Q. P., Yao, Y. G., Sun, C., Bandelt, H. J., Zhu, C. L. and Zhang, Y. P., 2003. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. *The American Journal of Human Genetics*, *73*(3): 671-676.

Krause, J., A. W.Briggs, M. Kircher, T. Maricic, N. Zwyns, A. Derevianko and S. Pääbo, 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. *Current Biology*, *20*(3): 231-236.

Krings, M., A. Stone, R. W. Schmitz, H. Krainitzki, M. Stoneking, S. Pääbo. 1997. Neandertal DNA sequences and the origin of modern humans. *Cell*. 90(1): 19-30.

Krings, M., H. Geisert, R. W. Schmitz, H. Krainitzki, S. Pääbo. 1999. DNA sequence of the mitochondrial hyperviable region II from the Neanderthal type specimen. Proc. *Natl. Acad. Sci. USA*. 96(10): 5581-85.

Kumar, S., R. R. Ravuri, P. Koneru, B. P. Urade, B. N. Sarkar, A. Chandrasekar, and V. R. Rao, 2009. Reconstructing Indian-Australian phylogenetic link. *BMC Evolutionary Biology*, *9*(1): 1.

Lacau, H., T. Gayden, M. Regueiro, S. Chennakrishnaiah, A. Bukhari, P. A. Underhill, Garcia- R. L. Bertrand, R. J. Herrera. 2012. Afghanistan from a Y-chromosome perspective. *Eur. J. Hum.Genet*., 20: 1063–1070.

Lafreniere, P. 2010. Adaptive Origins: Evolution and Human Development. Taylor & Francis.psycology press. p. 90.

Lahn, B. T., N. M. Pearson and K. Jegalian. 2001. The human Y chromosome, in the light of evolution. *National Review Genetics,* 2 (3): 207–16.

Lee, E. Y., K. J. Shin, A. Rakha, J. E. Sim, M. J. Park, N. Y. Kim, W. I. Yang and H. Y. Lee. 2014. Analysis of 22 Y chromosomal STR haplotypes and Y haplogroup distribution in Pathans of Pakistan. *Forensic Sci Int Genet*., 11: 111–116.

Lembring, M. 2013. Application of Mitochondrial DNA Analysis in Contemporary and Historical Samples.*Digital Comprehensive Summaries of Uppsala Dissertationsfrom the Faculty of Medicine,* 950: 62.

Lewis, M. P. 2009. Pashto, Northern. Ethnologue: Languages of the World, Sixteenth edition.SIL International.Dallas, Texas.

Liu, W., C. Z. Jin, Y. Q. Zhang, Y. Z. Cai, S. Xing, X. J. Wu, H. Cheng, R. L. Edwards, W. S. Pan, D. G. Qin and Z. S. An. 2010. Human remains from Zhirendong, South China and modern human emergence in East Asia. *Proceedings of the National Academy of Sciences* 107(45): 19201–19206.

Lott, M. T., J. N. Leipzig, O. Derbeneva, H. M. Xie, D. Chalkia, M. Sarmady, V. Procaccio, D. C. Wallace. 2013. mtDNA variation and analysis using MITOMAP and MITOMASTER. *Current Protocols in Bioinformatics*, 1(123): 1-26.

Luis, J.R., D. J. Rowold, M. Regueiro, *et al*., 2004. The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. *Am. J. Hum. Genet*., 74 (3): 532–44.

Maca-Meyer N., A. M. González, J. M. Larruga, C. Flores, V. M. Cabrera. 2001. Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet*., 2(1): 1.

Macaulay, V., C. Hill, A. Achilli, C. Rengo, D. Clarke, W. Meehan, J. Blackburn, O. Semino, R. Scozzari, F. Cruciani and A. Taha. 2005. A Response to Tracing Modern Human Origins. *Science*, *309*(5743): 1995b-1997b.

Malmström, H., M. T. P. Gilbert, M. G. Thomas, M. Brandström, J. Storå, P. Molnar, P. K. Andersen, C. Bendixen, G. Holmlund, A. Götherström and E. Willerslev, 2009. Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. *Current Biology*, *19*(20): 1758-1762.

Mansoor, A., K. Mazhar, S. Khaliq, A. Hameed, S. Rehman, S. Siddiqi, M. Papaioannou, L. L. Cavalli-Sforza, S. Q. Mehdi, Q. Ayub. *2004.* Investigation of the Greek ancestry of populations from northern Pakistan. *Hum. Genet.*, 114(5): 484–490.

Marisi, A., S. R. P. LINE. 2007. A simple and cost–effective protocol for DNA isolation from buccal epithelial cells.*Braz. Dent. J.*, 18(2):148-152.

Marshal, J. and J. P. Vogel, 1902-3. Excavations at Charsada in Frontier Province. *Archeological Survey of India Annual report*, 141-184.

Martinez, L., S. Mirabal, J. R. Luis and R. J. Herrera, 2008. Middle Eastern and European mtDNA lineages characterize populations from eastern Crete. *American journal of physical anthropology*, *137*(2): 213-223.

McElreavey, K. and L. Quintana-Murci. 2005. A population genetics perspective of the Indus Valley through uniparentally-inherited markers. *Ann. Hum. Biol.*, 32(32):154–162.

Mcrindle, J. W. 1992. The Invasion of India by Alexander the Great, New edition, New Delhi.

Mehdi, S. Q., R. Qamar, Q. Ayub, S. Kaliq, A. Mansoor, M. Ismail, M. F. Hammer, P. A. Underhill, L. L. Cavalli-Sforza. 1999. The origins of Pakistani populations:

evidence from Y chromosome markers. In: S. S. Papiha, R. Deka, R. Chakraborty (eds). Genomic diversity: applications in human population genetics. Kluwer Academic/Plenum Publishers, New York, pp 83–90.

Mendez, F. L., T. M. Karafet, T. Krahn, H. Ostrer, H. Soodyall and M. F. Hammer, 2011. Increased Resolution of Y Chromosome Haplogroup T Defines Relationships among Populations of the Near East, Europe, and Africa. *Human Biology*, 83(1): 39–53.

Mendizabal, I., C. Valente, A. Gusmão, C. Alves, V. Gomes, A. Goios, W. Parson, F. Calafell, L. Alvarez, A. Amorim and L.Gusmão, 2011. Reconstructing the Indian origin and dispersal of the European Roma: a maternal genetic perspective. *PloS one*, *6*(1): e15988.

Merriwether, D. A., A. G. Clark, S. W. Ballinger, T. G. Schurr, H. Soodyall, T. Jenkins, S. T. Sherry and D. C. Wallace. 1991. The structure of human mitochondrial DNA variation. *J. Mol. Evol.*, 33: 543-555.

Metspalu, M., T. Kivisild, E. Metspalu, J. Parik, G. Hudjashov, K. Kaldma, P. Serk, M. Karmin, D. M. Behar, M. T. Gilbert, P. Endicott, S. Mastana, S. S. Papiha, K. Skorecki, A. Torroni, R. Villems. 2004. Most of the extant mtDNA boundaries in South and Southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC genetics*, 5(1): 26.

Minorsky, V. 1940. The Turkish Dialect of the Khalaj, *Bulletin of the School of Oriental Studies.* University of London, 10(2): 417-437.

Mirabal, S., M. Regueiro, A. M. Cadenas, L. L. Cavalli-Sforza, P. A. Underhill, D. A. Verbenko, S. A. Limborska, and R. J. Herrera, 2009.Y-chromosome distribution within the geo-linguistic landscape of northwestern Russia. *European Journal of Human Genetics*, *17*(10): 1260-1273.

Mohammed, A. and H. Imad. 2013. Autosomal STR: From locus information to next generation sequencing technology. *Res. J. Biotechnol*., 8(10): 92-105.

Mohyuddin, A., Q. Ayub, R. Qamar, T. Zerjal, A. Helgason, S. Q. Mehdi and C. Tyler Smith, 2001. Y-chromosomal STR haplotypes in Pakistani populations. *Forensic science international*, *118*(2): 141-146.

Mosk, C., 2013. *Nationalism and Economic Development in Modern Eurasia* (Vol. 62). Routledge.

Muhammadzai, Q. J. 2012. From Pushkalavati to Charsada Via Hashtnagar. *Journal of Asian Civilizations*, 35(2): 144-62.

Murci, Q. L., R. Chaix, R. S. Wells, D. M. Behar, H. Sayar, R. Scozzari, C. Rengo, N. AlZahery, O. Semino, A. S. Santachiara-Benerecetti, A. Coppa, Q. Ayub, A. Mohyuddin, C. Tyler-Smith, S. Qasim Mehdi, A. Torroni, K. McElreavey. 2004. Where west meets east, the complex mtDNA landscape of the southwest and Central Asian corridor. *Am. J. Hum.Genet*., 74(5):827–45.

Murdock, G. P. 1967. Ethnographic Atlas: a summary. *Ethnology*, 6(2): 109-236.

Murray, J. W. 1899. A Dictionary of the Pathan Tribes on the North-West Frontier of

India. Calcutta: Office of the Superintendent, Government Printing, India. 157.

Nasidze, I., D. Quinque, I. Udina, S. Kunizheva and M. Stoneking, 2007. The Gagauz, a linguistic enclave, are not a genetic isolate. *Annals of human genetics*, *71*(3): 379-389.

Nasidze, I., D. Quinque, M. Rahmani, S. A. Alemohamad and M. Stoneking, 2008. Close Genetic Relationship Between Semitic speaking and Indo European speaking Groups in Iran. *Annals of human genetics*, *72*(2): 241-252.

Nasidze, I., D. Quinque, M. Rahmani, S. A. Alemohamad and M. Stoneking, 2006. Concomitant replacement of language and mtDNA in South Caspian populations of Iran. *Current biology*, *16*(7): 668-673.

Nasidze, I., D. Quinque, M. Rahmani, S. A. Alemohamad and M. Stoneking, 2008. Close Genetic Relationship between Semitic speaking and Indo European speaking Groups in Iran. *Annals of human genetics*, *72*(2): 241-252.

Nei, M. 1987. Molecular Evolutionary Genetics, Columbia University Press, New York.

Noor, S., S. Ali, M. Eaaswarkhanth and I. Haque. 2009. Allele frequency distribution for 15 autosomal STR loci in Afridi Pathan population of Uttar Pradesh, India. *Legal Medicine*, *11*(6): 308-311.

Oddie, G. A. 1991. Religion in South Asia: religious conversion and revival movements in South Asia in medieval and modern times. Columbia: South Asia Books.

Official Gazette of Tehsil/District Nowshera, 1870: 23–26.

Olivio, P. D., M. J. Van de Walle, P. J. Laipis, W. W. Hausewirth, 1983. Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop. *Nature*, 306: 400-402.

Ovchinnikow, I. V., A. Götherström, G. P. Romanowa, V. M. Kharitonov, K. Liden, W. Goodwin. 2000. Molecular analysis of Neanderthal DNA from the northern Caucasus. *Nature,* 404(6777): 490-493.

Palanichamy, M. G., B. Mitra, C. L. Zhang, M. Debnath, G. M. Li, H. W. Wang, S. Agrawal, T. K. Chaudhuri, and Y. P. Zhang, 2015. West Eurasian mtDNA lineages in India: an insight into the spread of the Dravidian language and the origins of the caste system. *Human genetics*, *134*(6): 637-647.

Palanichamy, M. G., B. Mitra, C. L. Zhang, M. Debnath, G. M. Li, H. W. Wang, S. Agrawal, T. K. Chaudhuri, and Y. P. Zhang, 2015. West Eurasian mtDNA lineages in India: an insight into the spread of the Dravidian language and the origins of the caste system. *Human genetics*, *134*(6): 637-647.

Passarino, G., G. L. Cavalleri, A. A. Lin, L. L. Cavalli-Sforza, A. L. Børresen-Dale and P. A. Underhill. 2002. Different genetic components in the Norwegian population revealed by the analysis of mtDNA and Y chromosome polymorphisms. Eur. J. Hum. Genet., 10(9): 521–529.

Peakall, R. and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research.*Molecular Ecology Notes,* 6(1): 288295.

Pesole, G., C. Gissi, A. De Chirico and C. Saccone. 1999. Nucleotide substitution rate of

mammalian mitochondrial genomes. *J. Mol. Evol*., 48(4): 429-434.

Piercy, R., K. M. Sullivan, N. Benson and P. Gill. 1993. The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *International journal of legal medicine*, 106(2): 85-90.

Purps, J.S., S. Siegert, S. Willuweit, M. Nagy, C. Alves, R. Salazar, S. M. Angustia, L. H. Santos, K. Anslinger, B. Bayer. 2014. A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Sci Int Genet*., 12:12-23.

Qamar, R., Q. Ayub, A. Mohyuddin, A. Helgason, K. Mazhar, A. Mansoor, T. Zerjal, C. Tyler-Smith, S.Q. Mehdi. 2002. Y-Chromosomal DNA Variation in Pakistan. *American Journal of Human Genetics*, 70(5): 1107–24.

Qamar, R., Q. Ayub, S. Khaliq, A. Mansoor, T. Karafet, S. Q. Mehdi, M. F. Hammer. 1999. African and Levantine origins of Pakistani YAP + Y chromosomes. *Hum Biol*., 71:745–755.

Quintana-Murci, L., O. Semino, H. J. Bandelt, G. Passarino, K. McElreavey, *et al*. 1999.Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. *Nature genetics*, 23(4): 437- 441.

Rajkumar, R., J. Banerjee, H. B. Gunturi, R. Trivedi and V. K. Kashyap. 2005. Phylogeny and antiquity of M macrohaplogroup inferred from complete mt DNA sequence of Indian specific lineages, *BMC Evolutionary Biology*, 5(1):26.

Rakha, A., K. J. Shin, J. A. Yoon, N. Y. Kim, M. H. Siddique, I. S. Yang, W. I. Yang and H. Y. Lee. 2011. Forensic and genetic characterization of mtDNA from Pathans of Pakistan. *Int J Legal Med*., 125(6):841–8.

Rawlinson, V. (trans.). 1858. Chapter 7 of *The History of Herodotus*, originally written 440 B.C.E.). Vol. 1.

Richards J. F. 1993. The New Cambridge History of India: The Mughal Empire. New York: Cambridge University Press. p. 50.

Richards, M., H-J. Bandelt, T. Kivisild and S. Oppenheimer. 2006. A Model for the Dispersal of Modern Humans out of Africa. In: H-J. Bandelt, V. Macaulay, M. Richards, editors. Human Mitochondrial DNA and the Evolution of Homo sapiens: Springer Berlin Heidelberg. pp. 225-265.

Rodriguez-Flores, J. L., K. Fakhro, F. Agosto-Perez, M. D. Ramstetter, L. Arbiza, T. L. Vincent, A. Robay, J. A. Malek, K. Suhre and L. Chouchane. 2016. Indigenous Arabs are descendants of the earliest split from ancient Eurasian populations. Genome Res. 26: 151–162.

Roewer, L., S. Willuweit, C. Krüger, M. Nagy, S. Rychkov, I. Morozowa, O. Naumova, Y. Schneider, O. Zhukova, M. Stoneking and I. Nasidze. 2008. Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. *International journal of legal medicine*, 122(3):219–223.

Roewer, L., S. Willuweit, M. Stoneking and I. Nasidze. 2009. A Y-STR database of Iranian

and Azerbaijanian minority populations. *Forensic Sci. Int. Genet*., 4(1): e53–e55.

Roostalu, U., I. Kutuev, E. Loogväli, E. Metspalu, K. Tambets, M. Reidla, E. Khusnutdinova, E. Usanga, T. Kivisild, R. Villems. 2006. Origin and Expansion of Haplogroup H, the Dominant Human Mitochondrial DNA Lineage in West Eurasia: The Near Eastern and Caucasian Perspective. *Mol Biol Evol*., 24(2): 436448.

Rose, H. A. 1997. A Glossary of the Tribes and Castes of the Punjab and North-West Frontier Province: L-Z. Vol. 3. Atlantic Publishers & Dist., p. 251.

Roy, O. 2006. Globalized Islam: The Search For A New Ummah. New York: Columbia University Press. P. 261.

Roy, J. N. and B. B. Kumar (ed). 2007. India and Central Asia: Classical to contemporary periods. Concept publishing company Dheli. P. 64.

Rozen, S., H. Skaletsky, J. D. Marszalek, P. J. Minx, H. S. Cordum, R. H. Waterston, R. K. Wilson & D. C. Page. 2003. Abundant gene conversion between arms of palindromes in human and ape Y-chromosomes. *Nature,* 423(6942): 873-876.

Runion, M. L. 2007. The history of Afghanistan. Connecticut: Greenwood Press.

Saccone, C., C. Gissi, C. Lanave, A. Larizza, G. Pesole, A. Reyas. 2000. Evolution of the mitochondrial genetic system: An overview. *Gene*, 261(1): 153-159.

Sahoo, S., A. Singh, G. Himabindu, J. Banerjee, T. Sitalaximi, S. Gaikwad, R. Trivedi, P. Endicott, T. Kivisild, M. Metspalu, R. Villems, V. K. Kashyap. 2006. A

prehistory of Indian Y chromosomes: Evaluating demic diffusion scenarios. *Proceedings of the National Academy of Sciences*, 103(4): 843–848.

Sanchez, J. J., M. Brion, W. Parson, A. J. Blanco-Verea, C. Borsting, M. Lareu, H. Niederstätter, H. Oberacher, N. Morling and A. Carracedo. 2004. Duplications of the Y-chromosome specific loci P25 and 92R7 and forensic implications. *Forensic Sci. Int*. 140(2):241–50.

Sánchez-Quinto, F., H. Schroeder, O.Ramirez, M. C. Ávila-Arcos, M. Pybus, I. Olalde, A. M. Velazquez, M. E. P. Marcos, J. M. V. Encinas, J. Bertranpetit and L.

Orlando, 2012. Genomic affinities of two 7,000-year-old Iberian hunter gatherers. *Current biology*, *22*(16): 1494-1499.

Semino, O., C. Magri, G. Benuzzi, A. A. Lin, N. Al-Zahery, V. Battaglia, L. Maccioni, C. Triantaphyllidis, P. Shen, P. J. Oefner and L. A. Zhivotovsky. 2004. Origin, Diffusion, and Differentiation of Y-Chromosome Haplogroups E and J: Inferences on the Neolithization of Europe and Later Migratory Events in the Mediterranean Area. *Am J Hum Genet.,* 74(5):1023-1034.

Semino, O., S. Benerecetti, A. Silvana, F. Francesco, L. L. Cavalli-Sforza and P. A. Underhill. 2002. Ethiopians and Khoisan Share the Deepest Clades of the Human Y-Chromosome Phylogeny. *Am J Hum Genet.*,70(1): 265–268.

Shan, W., A. Ablimit, W. Zhou, F. Zhang, Z. Ma and X. Zheng. 2014. Genetic

polymorphism of 17 Y chromosomal STRs in Kazakh and Uighur populations from Xinjiang, China. *International journal of legal medicine*, 128:743–744.

Sharma, S., E. Rai, A. K. Bhat, A. S. Bhanwer and R. N. Bamezai. 2007. A novel subgroup Q5 of human Y-chromosomal haplogroup Q-M242 in India. *BMC Evol. Biol.* **7** (1): 232.

Shen, G., W. Wang, Q. Wang, J. Zhao, K. Collerson, C. Zhou and P. V. Tobias.2002. USeries dating of Liujiang hominid site in Guangxi, Southern China. *Journal of Human Evolution*, 43(6): 817–829.

Shen, P., T. Lavi, T. Kivisild, V.Chou, D. Sengun, D. Gefel, I. Shpirer, E. Woolf, J. Hillel, M. W. Feldman, and P. J. Oefner. 2004. Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y Chromosome and mitochondrial DNA sequence Variation. *Human Mutation*, *24*(3): 248-260.

Shlush, L. I., G. Atzmon, R. Weisshof, D. Behar, G. Yudkovsky, N. Barzilai and K. Skorecki, 2008. Ashkenazi Jewish centenarians do not demonstrate enrichment in mitochondrial haplogroup J. *PloS one*, *3*(10): e3425.

Shou, W. H., E. F. Qiao, C. Y. Wei, Y. L. Dong, S. J. Tan, H. Shi, W. R. Tang and C. J. Xiao. 2010. Y- Chromosome Distributions among Populations in Northwest China Identify Significant Contribution from Central Asian Pastoralists and Lesser Influence of Western Eurasians. Journal of Human Genetics, 55(5): 314-322.

Siddiqi, M. H., T. Akhtar, A. Rakha, G. Abbas, A. Ali, N. Haider, A. Ali, S. Hayat, S.

Masooma, J. Ahmad and M. A. Tariq, 2015.Genetic characterization of the Makrani people of Pakistan from mitochondrial DNA control-region data. *Legal Medicine*, *17*(2): 134-139.

Sinclair, A. H., P. Berta, M. S. Palmer, J. R. Hawkins, B. L. Griffiths, M. J. Smith, J. W. Foster, A. M. Frischauf, R. Lovell-Badge and P. N. Goodfellow. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature*, 346: 240-244.

Skaletsky, H., T. Kuroda-Kawaguchi, P. J. Minx, H. S. Cordum, L. Hillier, L. G. Brown, S. Repping, *T.* Pyntikova, J. Ali, T. Bieri and A. Chinwalla. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature,* 423(6942): 825–837.

Soares, P., L. Ermini, N. Thomson, M. Mormina, T. Rito, A. Röhl, A. Salas, S. Oppenheimer, V. Macaulay and M. B. Richards. 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. *The American Journal of Human Genetics*, *84*(6): 740-759.

Stoljarova, M., J. L. King, M. Takahashi, A. Aaspõllu and B. Budowle. 2016. Whole mitochondrial genome genetic diversity in an Estonian population sample. *International journal of legal medicine*, *130*(1): 67-71.

Stoneking, M. 2000. Hyperviable sites in the mtDNA control region are mutational hotspots. *American Journal of Human Genet*ics, 67(4): 1029-1032.

Syed Saleem Shahzad. 2007. How the Taliban kept their coffers full. *Asia Times Online*.10.

Thanseem, I., K. Thangaraj, G. Chaubey, V. Singh, L. V. K. S. Bhaskar, B. M. Reddy, A. G. Reddy and L. Singh. 2006. Genetic affinities among the lower castes and tribal groups of India: Inference from Y chromosome and mitochondrial DNA. *BMC Genetics*, 7(1): 42.

Thompson, W. E., J. Ramalho-Santos and P. Sutovsky. 2003. Ubiquitination of prohibitin in mammalian sperm mitochondria: possible roles in the regulation of mitochondrial inheritance and sperm quality control. *Biology of reproduction,* 69(1): 254-260.

Thomson, R., J. K. Pritchard, P. Shen, P. J. Oefner and M. W. Feldman. 2000. Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc. Natl Acad. Sci. USA,* 97(13): 7360–7365.

Torroni, A., A.Achilli, V. Macaulay, M. Richards and H. J. Bandelt, 2006.Harvesting the fruit of the human mtDNA tree. *TRENDS in Genetics*, *22*(6): 339-345.

Torroni, A., K. Huoponen, P. Francalacci, M. Petrozzi, L. Morelli, *R.* Scozzari, D. Obinu,M. L. Savontaus, D.C. Wallace. 1996. Classification of European mtDNAs from an analysis of three European populations. *Genetics*, 144(4): 1835-1850.

Underhill, P. A., G. Passarino, A. A. Lin, P. Shen, M. Mirazón Lahr, R. A. Foley, P. J. Oefner and L. L. Cavalli-Sforza. 2001. The phylogeography of Y-chromosome binary haplotypes and the origin of modern human populations. *Annual Human Genetics*, 65(1): 43-62.

Underhill, P. A., L. Jin, A. A. Lin, S. Q. Mehdi, T. Jenkins, D. Vollrath, R.W. Davis, L. L. Cavalli-Sforza, P. J. Oefner. 1996. Detection of numerous Y-chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Research,*7(10):996-1005.

Underhill, P. A., N. M. Myres, S. Rootsi, M. Metspalu, L. A. Zhivotovsky, R. J. King, A. A. Lin, C. E. T. Chow, O. Semino, V. Battaglia and I.Kutuev, 2010. Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *European Journal of Human Genetics*, *18*(4): 479-484.

Van, O. M. and M. Kayser. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human mutation*, 30(2): E386-394.

Vermeulen, M., A. Wollstein, K. Van der Gaag, O. Lao, Y. Xue, Q. Wang, L. Roewer, H. Knoblauch, C. Tyler-Smith, P. de Knijff and M. Kayser. 2009. Improving global and regional resolution of male lineage differentiation by simple single-copy Ychromosomal short tandem repeat polymorphisms. *Forensic Sci. Int.*, 3(4): 205–213.

Vigilant, L., M. Stoneking, H. Harpending, K. Hawkes and A. C. Wilson. 1991. African populations and the evolution of human mitochondrial DNA. *Science,* 253(5027): 1503–7.

Walkinsha, M., L. Strickland, H. Hamilton, K. Denning and T. Gayley. 1996. DNA Profiling in two Alaskan Native Populations Using HLA-DQA1, PM, and D1S80 Loci. *J. Forensic Sci.*,41(3): 478-484.

Wallace, D. C., M. T. Lott, M. D. Brown, K. Huoponen, A. Toroni. 1995. Human gene

mapping 1995: A compendium. John Hopkins University Press, Baltimore.

Wells, R. S., 2007. Deep ancestry: inside the genography project. The landmark DNA quest to decipher our distant past. *National Geographics*, p. 103.

Wells, R. S., N. Yuldasheva, R. Ruzibakiev, P. A. Underhill, I. Evseeva, J. Blue-Smith, L. Jin, B. Su, R. Pitchappan, S. Shanmugalakshmi and K. Balakrishnan, 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proceedings of the National Academy of Sciences*, *98*(18): 10244-10249.

Whale, J. W., 2012.Mitochondrial DNA analysis of four ethnic groups of Afghanistan. *Portsmouth: University of Portsmouth*.

Wheeler, R. E. M. 1962. Charsadda: A Metropolis of the North-West frontier. Being a Report on the Excavations of 1958. Government of Pakistan and the British Academy. Oxford University PressLondon.

White, T. D., B. Asfaw, D. DeGusta, H. Gilbert, G.D. Richards, G. Suwa, F. C. Howell. 2003. Pleistocene Homo sapiens from Middle Awash, Ethiopia. *Nature,* 423(6491): 742–747.

Wilford, J. N. 2011. Fossil Teeth Put Humans in Europe Earlier Than Thought. *The New York Times.*

Wilson, M. R., J. DiZinno, D. Polanskey, J. Replogle and B. Budowle. 1995. Validation of mitochondrial DNA sequencing for forensic casework analysis. *International Journal of Legal Medicine*, 108(2): 68–74.

Wolpert, S. 2000. A new history of India. New York: Oxford University Press.

Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Research,* 12(2): 339-348.

Yan, S., C. C. Wang, H. Li, S. L. Li and L. Jin. The Genographic Consortium. 2011. An Updated Tree of Y-chromosome Haplogroup O and Revised Phylogenetic Positions of Mutations P164 and PK4. *European Journal of Human Genetics*, 19(9): 1013–1015.

Ye, Y., J. Gao, G. Fan, L. Liao and Y. Hou. 2015. Population genetics for 23 Y-STR loci in Tibetan in China and confirmation of DYS448 null allele. *Forensic Sci Int Genet.,* 16: e7–e10.

Zegura, S. L., T. M. Karafet, L. A. Zhivotovsky and M. F. Hammer. 2004. High-Resolution SNPs and Microsatellite Haplotypes Point to a Single, Recent Entry of Native American Y Chromosomes into the Americas. *Molecular Biology and Evolution* 21(1): 164–75.

Zhivotovsky, L. A., S. Ahmed, W. Wang and M. Bittles. 2001. The forensic DNA implications of genetic differentiation between endogamous communities. *Forensic Sci. Int.*, 119(3): 269-72.

**APPENDIX I**

**ETHNOGENETIC PROJECT DEPARTMENT OF GENEITCS HAZARA UNIVERSITY, MANSEHRA**
Tel: 0997-414131 Fax: 0997-530046
URL:http//www.hu.edu.pk

**NOTE:** This Performa is the property of the **ETHNO GENETIC PROJECT (HEC)** Department of Genetics Hazara University Mansehra Pakistan. This information will be used for research purpose only and will be kept secret.

**Ref. #:** _____

**Lab code:** _____

**Date:** _____

**DETAILS OF PARTICIPANT**

**Name:** _____ **Father's Name:** _____

**Age:** _____ **Sex:** _____ **Native Language:** _____

**Ethnic Group:** _____ **Caste** _____

**Collection point:** _____

**Home Address:** _____

**Biological Samples Collected**

■ Dental impressions ■ Saliva

**Consent of Participant**

~ The researchers have informed me about the purpose of sample collection and their use in human genetics research. I have provided samples voluntarily to improve science in Pakistan. **Signature of Participant:** _____ **APPENDEX II**

**STOCK REAGENTS**

**Phenol:Chloroform Mixture (1:1)**

For each sample 200uL of phenol and 200uL of chloroform were used.

**Lysis Buffer**

500mM Tris-base

250mM EDTA

5% SDS

Proteinase K 75ug/mL of lysis solution

B-mercaptoethanol (14.4M), 1uL/mL of lysis solution

**50X TAE buffer**

M Tris-HCl PH8

0.5 M EDTA

Make up to 1 L with $dH_2O$ and autoclave

**Bromophenol blue dye**

0.1 g bromophenol blue

Dissolve

Adjust volume to 100 ml with $dH_2O$, stir overnight

PH to 8.0

Filter through Whatmann filter paper

Store at room temperature

**10 mg/ml Ethidium bromide (EtBr)**

Add 1 g of Ethidium bromide to

100 ml of ddH2O

Stir for several hours until completely dissolved Store wrapped in aluminum foil at

4ÝC **1M Tris**

**HCl**

For 1L:

121.1g Tris

800ml m $H_2O$

~42ml HCL

Autoclave for 20 minutes on liquid setting.

**SDS 20%**

Dissolve 20 g SDS (sodium dodecyl sulfate or sodium lauryl sulfate) in $H_2O$ to 100ml total with stirring (it may be necessary to heat the solution slightlyto fully dissolve the powder).

Filters sterilize using a 0.45-□m filter.

Keep 10ml of 20% solution in a Falcon tube with liquids.

**dNTP**

Master stock: 100mM.

Create a stock solution that is 2.5mM for each dNTP.

Add 25ul of each dNTP to 900ul of $H_2O$ for a total stock solution size of 1ml. The final recommended final concentration for each dNTP in the PCR reaction is usually 200uM. Therefore, for a 25ul PCR reaction, add 2ul of the above 2.5mM stock solution.

**For Electrophoresis**

6X Loading dye Solution (from Fermentas)

10mM tris-HCl (pH7.6)

10mM EDTA

0.005% bromophenol blue

0.005% xylene cyanol FF and

10% glycerol

**TAE buffer, 50x**

242 g Tris base

57.1 ml glacial acetic acid

37.2 g Na2EDTA • $2H_2O$ (2mM)

$H_2O$ to 1 liter

This solution does not normally needs to be sterilized. The Tris base and acetic acid correspond to 40 mM Trisacetate

**TBE buffer, 10x**

108 g Tris base (890mM)

55g boric acid (890mM)

40 ml 0.5 M EDTA, pH 8.0 (20mM)

$H_2O$ to 1 liter

10x and 5x TBE tend to precipitate over time. If convenient, dilute to 2x or 1x immediately, or stir continuously. This solution does not normally need to be sterilized.

**SB buffer, 50X**

0.5M Sodium Hydroxide (NaOH), pH adjusted to 8.5 with boric acid ($H_3BO_{3)}$.

OR

250mM Disodium borate decahydrate ($Na_2B_4O_7 \cdot 10H_2O$)

DNA Ladder (Fermetas)

For:

**GeneRuler DNA Ladder Mix**

**GeneRuler 1kb DNA Ladder Plus**
Add 20 ul of DNA Ladder, 20

ul of Loading Dye Solution

80 ul of water.

**Primer Dilutions for PCR and Sequencing**

Dilute primers to 5ul/ml for PCR and 10 ul/ml for Sequencing.

Master stock, 100 µM

100 µM = X nmoles lyophilized primer + (X × 10 µl molecular grade $H_2O$)

To determine the amount of $H_2O$ to add to the lyophilized primer multiply the number of nmol of primer in the tube by 10 and that will be the amount of $H_2O$ to add to make a 100 µM primer stock.

The original primer tubes are often used for this 100 µM stock.

Master stock primers newly suspended in $H_2O$ should be allowed to sit at room temperature for 10 minutes before they are used for working stock dilutions. Mix well before making working stock dilutions.

**Working stock, 10 µM, 5 µM**

Dilute the primer master stock in a sterile micro centrifuge tube 1:10 with molecular grade $H_2O$.

**Dilution Calculation Example**

Final Volume x Final Concentration = Starting Concentration x X

Example: 100 ml x 5ug/ml = (x ml) x 100 ug/ml

$x = 5$ ml

so add 5 ml of stock and 95 ml $H_2O$

**Annexure-**

**1** List of 23 Y-STR Haplotypes and their respective haplogroups in five Pathan populations of Charsada and Mardan district.

| | DYS576 | DYS389I | DYS448 | DYS389II | DYS19 | DYS391 | DYS481 | DYS549 | DYS533 | DYS438 | DYS437 | DYS570 | DYS635 | DYS390 | DYS439 | DYS392 | DYS643 | DYS393 | DYS458 | DYS385a | DYS385b | DYS456 | YGATAH4 | Haplogroup |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MM01 | 20 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MM02 | 16 | 14 | 21 | 31 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| MM03 | 17 | 13 | 19 | 29 | 14 | 11 | 21 | 13 | 13 | 12 | 15 | 17 | 23 | 24 | 13 | 13 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1b |
| MM04 | 18 | 12 | 19 | 29 | 14 | 10 | 23 | 13 | 12 | 10 | 15 | 16 | 23 | 21 | 11 | 16 | 10 | 11 | 15 | 14 | 15 | 16 | 12 | L |
| MM05 | 18 | 13 | 20 | 29 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 17 | 25 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MM06 | 17 | 13 | 20 | 29 | 16 | 10 | 23 | 13 | 12 | 11 | 14 | 21 | 23 | 25 | 12 | 11 | 10 | 13 | 18 | 12 | 14 | 16 | 13 | R1a |
| MM07 | 16 | 13 | 19 | 29 | 13 | 10 | 26 | 11 | 11 | 11 | 15 | 18 | 24 | 22 | 11 | 15 | 12 | 13 | 17 | 11 | 15 | 16 | 10 | Q |
| MM08 | 20 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 18 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MM09 | 15 | 12 | 19 | 28 | 14 | 10 | 23 | 13 | 12 | 11 | 15 | 14 | 23 | 22 | 14 | 14 | 10 | 11 | 16 | 13 | 17 | 16 | 12 | L |
| MM10 | 16 | 13 | 21 | 30 | 16 | 11 | 18 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 20 | 13 | 16 | 16 | 12 | G2a |
| MM11 | 20 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM12 | 17 | 13 | 19 | 30 | 15 | 10 | 24 | 13 | 12 | 10 | 16 | 16 | 21 | 22 | 13 | 14 | 12 | 12 | 17 | 9 | 15 | 15 | 11 | L |
| MM13 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM14 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM15 | 19 | 13 | 20 | 30 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 11 | R1a |
| MM16 | 20 | 14 | 20 | 31 | 16 | 10 | 23 | 12 | 12 | 11 | 14 | 17 | 23 | 24 | 12 | 11 | 10 | 13 | 17 | 11 | 16 | 16 | 13 | R1a |
| MM17 | 16 | 13 | 19 | 29 | 13 | 10 | 26 | 11 | 11 | 11 | 15 | 18 | 24 | 22 | 11 | 15 | 12 | 13 | 17 | 11 | 15 | 16 | 10 | Q |
| MM18 | 19 | 13 | 20 | 30 | 15 | 11 | 25 | 13 | 12 | 11 | 14 | 18 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM19 | 18 | 12 | 19 | 29 | 14 | 10 | 23 | 13 | 12 | 10 | 15 | 16 | 23 | 21 | 11 | 16 | 10 | 11 | 15 | 14 | 16 | 16 | 11 | L |
| MM20 | 17 | 13 | 20 | 30 | 16 | 11 | 24 | 12 | 12 | 11 | 14 | 21 | 23 | 25 | 10 | 11 | 10 | 13 | 17 | 11 | 15 | 15 | 12 | R1a |
| MM21 | 16 | 13 | 21 | 30 | 15 | 11 | 19 | 14 | 10 | 10 | 16 | 19 | 24 | 23 | 11 | 11 | 12 | 13 | 20 | 13 | 16 | 16 | 12 | G2a |
| MM22 | 19 | 12 | 19 | 28 | 14 | 10 | 24 | 1 | 10 | 11 | 16 | 21 | 26 | 23 | 10 | 10 | 9 | 13 | 17 | 11 | 18 | 15 | 12 | Q |
| MM23 | 17 | 12 | 19 | 29 | 15 | 11 | 24 | 12 | 9 | 9 | 15 | 15 | 21 | 24 | 11 | 11 | 9 | 13 | 16 | 14 | 18 | 13 | 11 | J2b |
| MM24 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 23 | 23 | 11 | 11 | 12 | 13 | 18 | 13 | 16 | 16 | 12 | G2a |
| MM25 | 18 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MM26 | 18 | 13 | 20 | 31 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM27 | 17 | 13 | 20 | 30 | 16 | 10 | 23 | 11 | 12 | 11 | 14 | 18 | 24 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 16 | 13 | R1a |
| MM28 | 15 | 12 | 19 | 28 | 15 | 10 | 23 | 13 | 11 | 9 | 15 | 13 | 22 | 24 | 12 | 11 | 10 | 13 | 16 | 13 | 18 | 13 | 11 | J2b |
| MM29 | 19 | 13 | 21 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 14 | 15 | 11 | 14 | 15 | 13 | R1a |
| MM30 | 19 | 13 | 20 | 29 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MM31 | 17 | 12 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 13 | 17 | 11 | 14 | 17 | 12 | R1a |
| MM32 | 18 | 13 | 20 | 31 | 16 | 10 | 24 | 11 | 12 | 11 | 14 | 20 | 24 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM33 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 11 | 13 | 15 | 11 | 14 | 14 | 14 | R1a |
| MM34 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| MM35 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 12 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| MM36 | 19 | 13 | 20 | 30 | 16 | 11 | 22 | 13 | 12 | 11 | 14 | 19 | 24 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MM37 | 16 | 13 | 19 | 28 | 14 | 11 | 22 | 13 | 12 | 12 | 15 | 17 | 24 | 24 | 13 | 14 | 10 | 12 | 18 | 11 | 14 | 16 | 13 | R1b |
| MM38 | 18 | 13 | 20 | 31 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM39 | 16 | 13 | 19 | 28 | 14 | 11 | 22 | 13 | 12 | 12 | 15 | 17 | 24 | 24 | 13 | 14 | 10 | 12 | 18 | 11 | 14 | 16 | 13 | R1b |
| MM40 | 18 | 13 | 20 | 31 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM41 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM42 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 18 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM43 | 18 | 13 | 20 | 31 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MM44 | 16 | 13 | 21 | 30 | 15 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 18 | 13 | 16 | 16 | 12 | G2a |
| MM45 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 15 | 15 | 12 | R1a |
| MM46 | 16 | 13 | 21 | 30 | 17 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 20 | 13 | 16 | 16 | 12 | G2a |
| MM47 | 19 | 13 | 20 | 30 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MM48 | 18 | 13 | 18 | 29 | 15 | 10 | 25 | 11 | 12 | 9 | 14 | 17 | 21 | 24 | 11 | 11 | 9 | 12 | 16 | 14 | 17 | 15 | 11 | H |
| MM49 | 18 | 13 | 20 | 31 | 15 | 11 | 23 | 11 | 13 | 9 | 14 | 20 | 23 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 16 | 12 | R1a |
| MM50 | 15 | 12 | 19 | 28 | 15 | 10 | 23 | 13 | 11 | 9 | 15 | 13 | 22 | 24 | 12 | 11 | 10 | 13 | 16 | 13 | 18 | 13 | 11 | J2b |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MM51 | 15 | 14 | 20 | 32 | 15 | 10 | 23 | 12 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 16 | 12 | 14 | 15 | 13 | R1a |
| MM52 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 16 | 13 | R1a |
| MM53 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 14 | 15 | 11 | 14 | 15 | 13 | R1a |
| MM54 | 21 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM55 | 16 | 13 | 21 | 30 | 16 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 22 | 13 | 16 | 16 | 12 | G2a |
| MM56 | 16 | 13 | 21 | 30 | 16 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 20 | 13 | 16 | 16 | 12 | G2a |
| MM57 | 18 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM58 | 20 | 13 | 20 | 31 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM59 | 19 | 14 | 21 | 29 | 14 | 10 | 24 | 11 | 12 | 9 | 15 | 16 | 23 | 23 | 12 | 11 | 10 | 13 | 17 | 13 | 16 | 15 | 11 | I2a |
| MM60 | 19 | 13 | 20 | 30 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MM61 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MM62 | 17 | 13 | 20 | 29 | 16 | 10 | 23 | 13 | 12 | 11 | 14 | 21 | 23 | 25 | 12 | 11 | 10 | 13 | 18 | 12 | 14 | 16 | 13 | R1a |
| MM63 | 19 | 14 | 20 | 30 | 15 | 10 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 13 | R1a |
| MM64 | 16 | 13 | 21 | 30 | 16 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 20 | 13 | 16 | 15 | 11 | G2a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MM65 | 18 | 12 | 19 | 29 | 14 | 10 | 23 | 13 | 12 | 10 | 15 | 16 | 23 | 21 | 11 | 16 | 10 | 11 | 15 | 14 | 16 | 16 | 12 | L |
| MM66 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MM67 | 17 | 13 | 20 | 31 | 14 | 10 | 26 | 12 | 11 | 8 | 14 | 15 | 22 | 23 | 11 | 11 | 10 | 14 | 16 | 16 | 20 | 15 | 12 | E1b1b |
| MM68 | 20 | 13 | 20 | 30 | 16 | 11 | 25 | 13 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MM69 | 17 | 13 | 19 | 29 | 14 | 11 | 21 | 13 | 13 | 12 | 15 | 17 | 23 | 24 | 12 | 13 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1b |
| MM70 | 17 | 13 | 19 | 30 | 13 | 10 | 22 | 12 | 11 | 11 | 16 | 15 | 25 | 23 | 10 | 10 | 10 | 13 | 16 | 13 | 19 | 15 | 12 | L |
| MM71 | 19 | 13 | 20 | 30 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MM72 | 19 | 14 | 21 | 29 | 14 | 10 | 24 | 11 | 12 | 9 | 15 | 16 | 23 | 23 | 12 | 11 | 10 | 13 | 17 | 13 | 16 | 15 | 11 | I2a |
| MM73 | 17 | 13 | 19 | 29 | 14 | 11 | 21 | 13 | 13 | 12 | 15 | 17 | 23 | 24 | 13 | 13 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1b |
| MM74 | 15 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 16 | 16 | 21 | 22 | 13 | 14 | 12 | 12 | 17 | 9 | 17 | 14 | 11 | L |
| MM75 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 16 | 16 | 21 | 22 | 13 | 14 | 12 | 12 | 16 | 9 | 16 | 15 | 11 | L |
| MZ01 | 18 | 14 | 20 | 30 | 15 | 10 | 23 | 12 | 12 | 9 | 14 | 17 | 20 | 22 | 12 | 11 | 9 | 12 | 18 | 14 | 19 | 16 | 12 | H |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MZ02 | 19 | 14 | 20 | 31 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 12 | 14 | 15 | 13 | R1a |
| MZ03 | 18 | 14 | 19 | 31 | 14 | 10 | 22 | 13 | 12 | 19 | 14 | 17 | 21 | 24 | 12 | 12 | 10 | 13 | 17 | 13 | 16 | 17 | 11 | E1b1b |
| MZ04 | 19 | 13 | 20 | 29 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ05 | 19 | 13 | 21 | 29 | 16 | 11 | 26 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ06 | 19 | 14 | 20 | 31 | 15 | 10 | 22 | 12 | 11 | 9 | 15 | 15 | 22 | 23 | 11 | 11 | 10 | 12 | 15 | 14 | 19 | 15 | 12 | L |
| MZ07 | 20 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MZ08 | 15 | 12 | 21 | 28 | 15 | 10 | 22 | 12 | 9 | 10 | 16 | 16 | 19 | 22 | 12 | 10 | 11 | 14 | 18 | 15 | 16 | 15 | 11 | G2a |
| MZ09 | 19 | 13 | 20 | 30 | 15 | 11 | 18 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ10 | 15 | 13 | 19 | 29 | 15 | 10 | 24 | 14 | 12 | 10 | 16 | 16 | 21 | 22 | 13 | 14 | 12 | 12 | 17 | 9 | 16 | 15 | 11 | L |
| MZ11 | 16 | 13 | 21 | 30 | 16 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| MZ12 | 20 | 13 | 20 | 30 | 16 | 11 | 25 | 13 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ13 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 15 | 16 | 12 | G2a |
| MZ15 | 19 | 13 | 20 | 30 | 16 | 10 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ14 | 14 | 13 | 19 | 29 | 14 | 10 | 23 | 13 | 12 | 11 | 16 | 17 | 25 | 23 | 11 | 10 | 11 | 13 | 17 | 12 | 17 | 16 | 11 | L |
| MZ16 | 14 | 13 | 19 | 29 | 14 | 10 | 23 | 13 | 12 | 11 | 16 | 17 | 25 | 23 | 11 | 10 | 11 | 13 | 17 | 12 | 17 | 16 | 11 | L |
| MZ17 | 18 | 13 | 19 | 27 | 14 | 10 | 26 | 13 | 10 | 11 | 15 | 18 | 23 | 23 | 11 | 12 | 10 | 13 | 17 | 13 | 21 | 15 | 11 | Q |
| MZ18 | 17 | 13 | 20 | 30 | 16 | 11 | 21 | 12 | 12 | 11 | 14 | 19 | 23 | 25 | 11 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MZ19 | 20 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 23 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ20 | 18 | 13 | 20 | 30 | 15 | 10 | 23 | 12 | 12 | 9 | 14 | 17 | 20 | 22 | 12 | 11 | 9 | 12 | 18 | 14 | 19 | 16 | 12 | H |
| MZ21 | 19 | 14 | 20 | 31 | 15 | 10 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | ? | 10 | 13 | 16 | 11 | 15 | 15 | 12 | R1a |
| MZ22 | 19 | 12 | 19 | 28 | 15 | 10 | 23 | 13 | 11 | 10 | 15 | 18 | 21 | 23 | 11 | 13 | 11 | 12 | 16 | 15 | 19 | 14 | 13 | L |
| MZ23 | 17 | 13 | 19 | 31 | 16 | 9 | 22 | 13 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MZ24 | 19 | 12 | 20 | 28 | 16 | 10 | 23 | 13 | 11 | 10 | 15 | 18 | 21 | 23 | 11 | 13 | 11 | 12 | 16 | 15 | 19 | 14 | 13 | L |
| MZ25 | 18 | 13 | 21 | 30 | 16 | 11 | 23 | 12 | 11 | 11 | 14 | 20 | 23 | 26 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 16 | 13 | R1a |
| MZ26 | 16 | 13 | 19 | 30 | 15 | 10 | 23 | 13 | 12 | 10 | 14 | 19 | 20 | 21 | 11 | 11 | 19 | 12 | 17 | 14 | 18 | 15 | 11 | H |
| MZ27 | 16 | 13 | 19 | 30 | 15 | 10 | 21 | 11 | 12 | 10 | 16 | 19 | 21 | 22 | 13 | 11 | 13 | 14 | 17 | 14 | 15 | 17 | 13 | G2a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MZ28 | 16 | 12 | 18 | 27 | 15 | 10 | 23 | 12 | 12 | 10 | 15 | 15 | 22 | 22 | 12 | 14 | 10 | 11 | 16 | 13 | 20 | 16 | 13 | L |
| MZ29 | 17 | 14 | 20 | 31 | 17 | 10 | 23 | 12 | 11 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 19 | 13 | 16 | 11 | 15 | 15 | 13 | R1a |
| MZ30 | 20 | 13 | 20 | 30 | 17 | 11 | 23 | 12 | 12 | 10 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MZ31 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| MZ32 | 20 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ33 | 19 | 13 | 19 | 28 | 15 | 10 | 24 | 12 | 11 | 11 | 16 | 18 | 24 | 23 | 11 | 10 | 10 | 14 | 17 | 12 | 20 | 16 | 11 | L |
| MZ34 | 20 | 14 | 19 | 30 | 13 | 10 | 25 | 12 | 11 | 11 | 16 | 17 | 23 | 21 | 11 | 15 | 11 | 13 | 17 | 12 | 17 | 15 | 10 | Q |
| MZ35 | 19 | 13 | 20 | 29 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ36 | 19 | 13 | 20 | 30 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MZ37 | 20 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ38 | 19 | 13 | 20 | 31 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ39 | 20 | 13 | 20 | 29 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 11 | R1a |
| MZ40 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MZ41 | 19 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ42 | 18 | 14 | 19 | 29 | 15 | 11 | 24 | 12 | 11 | 11 | 17 | 17 | 24 | 23 | 11 | 10 | 10 | 14 | 17 | 13 | 20 | 16 | 11 | L |
| MZ43 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 15 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ44 | 19 | 13 | 20 | 30 | 17 | 11 | 24 | 12 | 12 | 11 | 14 | 19 | 24 | 25 | 10 | 11 | 11 | 13 | 15 | 11 | 14 | 15 | 14 | R1a |
| MZ45 | 21 | 14 | 20 | 31 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ46 | 16 | 13 | 19 | 29 | 15 | 10 | 23 | 12 | 12 | 10 | 14 | 18 | 21 | 22 | 12 | 11 | 9 | 12 | 18 | 16 | 17 | 15 | 12 | H |
| MZ47 | 18 | 13 | 19 | 28 | 15 | 10 | 24 | 12 | 11 | 11 | 16 | 18 | 24 | 23 | 11 | 10 | 10 | 14 | 17 | 13 | 20 | 16 | 11 | L |
| MZ49 | 18 | 13 | 20 | 30 | 16 | 10 | 24 | 13 | 12 | 11 | 14 | 18 | 24 | 25 | 10 | 11 | 8 | 13 | 16 | 11 | 14 | 16 | 13 | R1a |
| MZ48 | 20 | 13 | 20 | 29 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 11 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ50 | 17 | 13 | 20 | 30 | 17 | 11 | 23 | 11 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 13 | R1a |
| MZ51 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 12 | 11 | 11 | 14 | 17 | 25 | 22 | 11 | 10 | 10 | 14 | 18 | 14 | 19 | 15 | 12 | Q |
| MZ52 | 19 | 13 | 20 | 30 | 15 | 10 | 26 | 12 | 12 | 11 | 14 | 18 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ53 | 15 | 13 | 20 | 30 | 15 | 7 | 21 | 13 | 12 | 8 | 14 | 11 | 20 | 21 | 12 | 11 | 10 | 12 | 14 | 13 | 16 | 15 | 11 | G2a |
| MZ54 | 15 | 13 | 20 | 30 | 15 | 7 | 25 | 13 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MZ55 | 16 | 13 | 19 | 30 | 15 | 7 | 24 | 13 | 12 | 10 | 16 | 16 | 21 | 22 | 13 | 14 | 12 | 12 | 17 | 9 | 16 | 15 | 11 | L |
| MZ56 | 19 | 13 | 20 | 29 | 16 | 11 | 22 | 12 | 12 | 11 | 14 | 21 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 14 | 13 | R1a |
| MZ57 | 18 | 14 | 20 | 30 | 16 | 9 | 28 | 12 | 13 | 10 | 14 | 16 | 21 | 23 | 12 | 12 | 10 | 14 | 19 | 13 | 15 | 15 | 12 | I2a (xI2a1) |
| MZ58 | 20 | 13 | 20 | 29 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ59 | 15 | 13 | 19 | 29 | 15 | 10 | 25 | 14 | 12 | 10 | 16 | 16 | 21 | 22 | 11 | 12 | 10 | 12 | 17 | 9 | 17 | 15 | 12 | L |
| MZ60 | 16 | 13 | 21 | 30 | 16 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| MZ61 | 16 | 13 | 19 | 30 | 16 | 10 | 23 | 13 | 12 | 10 | 14 | 19 | 20 | 22 | 11 | 11 | 9 | 12 | 17 | 14 | 18 | 15 | 11 | H |
| MZ62 | 19 | 13 | 20 | 30 | 16 | 11 | 24 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ63 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MZ64 | 18 | 14 | 20 | 31 | 16 | 10 | 22 | 12 | 11 | 9 | 15 | 15 | 22 | 23 | 11 | 11 | 10 | 12 | 15 | 14 | 20 | 15 | 12 | L |
| MZ65 | 19 | 12 | 20 | 28 | 15 | 10 | 27 | 12 | 11 | 11 | 14 | 20 | 24 | 24 | 12 | 11 | 11 | 12 | 19 | 12 | 13 | 15 | 11 | Q |
| MZ66 | 20 | 13 | 20 | 30 | 15 | 12 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 15 | 15 | 12 | R1a |
| MZ67 | 16 | 12 | 19 | 28 | 15 | 11 | 23 | 12 | 9 | 9 | 15 | 15 | 21 | 24 | 11 | 11 | 9 | 13 | 16 | 13 | 19 | 13 | 11 | J2b |
| MZ68 | 18 | 10 | 22 | 27 | 13 | 10 | 23 | 12 | 11 | 10 | 14 | 14 | 21 | 24 | 14 | 11 | 12 | 13 | 17 | 16 | 17 | 16 | 11 | E1b1b |
| MZ69 | 17 | 13 | 19 | 29 | 15 | 10 | 21 | 13 | 12 | 8 | 14 | 19 | 20 | 21 | 12 | 12 | 9 | 12 | 14 | 16 | 18 | 15 | 11 | H |
| MZ70 | 18 | 13 | 20 | 30 | 16 | 10 | 23 | 12 | 11 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 14 | 16 | 11 | 14 | 15 | 12 | R1a |
| MZ71 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 3 | 18 | 13 | 17 | 16 | 12 | G2a |
| MZ72 | 19 | 13 | 20 | 32 | 15 | 10 | 23 | 12 | 13 | 11 | 14 | 18 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MZ73 | 18 | 13 | 19 | 29 | 13 | 10 | 25 | 14 | 11 | 11 | 16 | 17 | 25 | 23 | 11 | 10 | 9 | 13 | 18 | 13 | 18 | 15 | 12 | Q |
| MZ74 | 18 | 13 | 19 | 29 | 13 | 10 | 25 | 14 | 11 | 11 | 16 | 17 | 25 | 23 | 11 | 10 | 9 | 13 | 18 | 13 | 18 | 15 | 12 | Q |
| MZ75 | 21 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| MD01 | 19 | 13 | 20 | 30 | 16 | 11 | 26 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MD02 | 21 | 11 | 17 | 26 | 14 | 10 | 27 | 10 | 11 | 10 | 14 | 21 | 22 | 24 | 8 | 11 | 11 | 12 | 19 | 12 | 18 | 16 | 11 | J1 |
| MD03 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 11 | 11 | 10 | 16 | 16 | 23 | 22 | 12 | 14 | 12 | 12 | 18 | 7 | 15 | 13 | 11 | L |
| MD04 | 18 | 14 | 20 | 31 | 16 | 10 | 22 | 12 | 11 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 13 | R1a |
| MD05 | 17 | 14 | 19 | 30 | 15 | 10 | 24 | 12 | 12 | 11 | 16 | 19 | 26 | 24 | 11 | 10 | 10 | 14 | 16 | 12 | 13 | 16 | 12 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MD06 | 16 | 13 | 21 | 30 | 16 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 14 | 12 | G2a |
| MD07 | 19 | 13 | 20 | 30 | 15 | 12 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD08 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MD09 | 20 | 11 | 17 | 26 | 14 | 10 | 27 | 10 | 11 | 10 | 14 | 21 | 22 | 24 | 8 | 11 | 11 | 12 | 19 | 12 | 16 | 16 | 11 | J1 |
| MD10 | 19 | 11 | 17 | 26 | 14 | 10 | 27 | 10 | 11 | 10 | 14 | 21 | 22 | 24 | 8 | 11 | 11 | 12 | 19 | 12 | 17 | 16 | 11 | J1 |
| MD11 | 19 | 13 | 20 | 31 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 14 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD12 | 20 | 11 | 17 | 26 | 0 | 10 | 27 | 10 | 11 | 10 | 14 | 21 | 22 | 24 | 8 | 11 | 11 | 12 | 19 | 12 | 17 | 16 | 11 | J1 |
| MD13 | 20 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MD14 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD15 | 20 | 11 | 17 | 26 | 14 | 10 | 27 | 10 | 11 | 10 | 14 | 21 | 22 | 24 | 8 | 11 | 11 | 11 | 19 | 12 | 17 | 16 | 11 | Q |
| MD16 | 15 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 16 | 16 | 21 | 22 | 13 | 14 | 12 | 12 | 17 | 9 | 16 | 15 | 11 | L |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MD17 | 18 | 13 | 20 | 30 | 17 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD18 | 15 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |
| MD19 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD20 | 19 | 13 | 20 | 30 | 15 | 10 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD21 | 18 | 12 | 21 | 32 | 14 | 10 | 23 | 12 | 11 | 10 | 16 | 19 | 22 | 22 | 11 | 11 | 9 | 13 | 15 | 14 | 16 | 15 | 12 | G2a |
| MD22 | 19 | 14 | 18 | 30 | 14 | 10 | 24 | 12 | 12 | 12 | 16 | 17 | 27 | 24 | 11 | 12 | 10 | 13 | 14 | 12 | 13 | 15 | 12 | R1b |
| MD23 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |
| MD24 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 11 | 11 | 10 | 16 | 16 | 22 | 22 | 12 | 14 | 12 | 12 | 18 | 7 | 15 | 13 | 11 | L |
| MD25 | 15 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |
| MD26 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 23 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD27 | 17 | 14 | 19 | 30 | 13 | 10 | 23 | 14 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 13 | 12 | 13 | 16 | 15 | 16 | 15 | 12 | Q |
| MD28 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 18 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD29 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD30 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 23 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD31 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 11 | 11 | 10 | 16 | 16 | 22 | 22 | 12 | 14 | 12 | 12 | 18 | 7 | 15 | 13 | 11 | L |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MD32 | 15 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |
| MD33 | 20 | 13 | 18 | 29 | 14 | 10 | 22 | 14 | 12 | 12 | 15 | 15 | 23 | 24 | 13 | 13 | 10 | 12 | 16 | 11 | 14 | 15 | 12 | R1b |
| MD34 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |
| MD35 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 14 | 12 | R1a |
| MD36 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |
| MD37 | 16 | 14 | 19 | 30 | 15 | 10 | 23 | 15 | 12 | 9 | 14 | 16 | 21 | 22 | 11 | 11 | 9 | 12 | 18 | 15 | 17 | 15 | 12 | H |
| MD38 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MD39 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD40 | 19 | 15 | 19 | 30 | 14 | 10 | 25 | 11 | 12 | 11 | 14 | 19 | 27 | 23 | 12 | 10 | 11 | 14 | 17 | 13 | 20 | 16 | 11 | Q |
| MD41 | 19 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD42 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 11 | 11 | 10 | 16 | 16 | 22 | 22 | 12 | 14 | 12 | 12 | 18 | 7 | 15 | 13 | 11 | L |
| MD43 | 17 | 13 | 21 | 30 | 16 | 12 | 23 | 12 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MD44 | 17 | 13 | 21 | 30 | 16 | 12 | 23 | 12 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| MD45 | 19 | 13 | 20 | 30 | 14 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 10 | 14 | 15 | 12 | R1a |
| MD46 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 16 | 13 | R1a |
| MD47 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |
| MD48 | 19 | 14 | 18 | 30 | 14 | 10 | 24 | 12 | 12 | 12 | 16 | 17 | 27 | 24 | 11 | 12 | 10 | 13 | 14 | 12 | 13 | 15 | 12 | R1b |
| MD49 | 18 | 12 | 19 | 28 | 15 | 10 | 22 | 12 | 11 | 10 | 14 | 16 | 19 | 23 | 13 | 12 | 11 | 12 | 16 | 11 | 16 | 14 | 12 | L |
| MD50 | 18 | 14 | 19 | 31 | 16 | 11 | 23 | 12 | 13 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 13 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MD51 | 19 | 13 | 19 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 23 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD52 | 19 | 13 | 20 | 30 | 16 | 11 | 26 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 16 | 12 | R1a |
| MD53 | 18 | 13 | 19 | 29 | 14 | 10 | 22 | 10 | 12 | 11 | 15 | 17 | 23 | 22 | 13 | 15 | 12 | 13 | 20 | 12 | 16 | 15 | 10 | L |
| MD54 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 20 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| MD55 | 15 | 13 | 19 | 31 | 15 | 10 | 26 | 13 | 11 | 11 | 16 | 19 | 26 | 24 | 11 | 10 | 10 | 14 | 16 | 12 | 18 | 16 | 12 | E1b1b |
| MD56 | 15 | 13 | 19 | 29 | 15 | 10 | 24 | 14 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |
| MD57 | 15 | 13 | 19 | 29 | 15 | 10 | 24 | 13 | 12 | 10 | 15 | 16 | 21 | 22 | 12 | 14 | 12 | 12 | 17 | 9 | 17 | 16 | 11 | L |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MD58 | 19 | 14 | 19 | 31 | 14 | 10 | 25 | 12 | 11 | 9 | 15 | 18 | 21 | 24 | 12 | 11 | 9 | 12 | 17 | 13 | 16 | 16 | 11 | J1 |
| MD59 | 19 | 14 | 19 | 30 | 14 | 10 | 25 | 13 | 11 | 11 | 15 | 19 | 25 | 23 | 10 | 10 | 10 | 14 | 16 | 13 | 22 | 15 | 11 | Q |
| MD60 | 18 | 13 | 20 | 30 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 10 | 14 | 15 | 12 | R1a |
| MD61 | 19 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 13 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD62 | 20 | 11 | 17 | 26 | 14 | 10 | 27 | 10 | 11 | 10 | 14 | 20 | 22 | 24 | 8 | 11 | 11 | 12 | 19 | 12 | 17 | 16 | 11 | J1 |
| MD63 | 16 | 14 | 21 | 31 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 16 | 24 | 23 | 11 | 11 | 12 | 13 | 22 | 13 | 16 | 15 | 12 | G2a |
| MD64 | 19 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 18 | 23 | 23 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD65 | 16 | 13 | 21 | 30 | 17 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 22 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| MD66 | 15 | 12 | 19 | 27 | 15 | 10 | 23 | 14 | 10 | 9 | 15 | 19 | 21 | 23 | 11 | 11 | 9 | 12 | 17 | 12 | 17 | 13 | 11 | J2b |
| MD67 | 19 | 13 | 20 | 29 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 17 | 11 | 14 | 15 | 12 | R1a |
| MD68 | 16 | 13 | 21 | 30 | 15 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 23 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 15 | 12 | G2a |
| MD69 | 15 | 13 | 20 | 30 | 15 | 10 | 24 | 11 | 12 | 10 | 16 | 15 | 13 | 22 | 12 | 14 | 14 | 12 | 17 | 9 | 16 | 15 | 11 | L |
| MD70 | 19 | 13 | 20 | 31 | 16 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD71 | 18 | 12 | 19 | 28 | 17 | 10 | 25 | 11 | 11 | 10 | 15 | 16 | 20 | 23 | 12 | 11 | 10 | 12 | 18 | 14 | 18 | 14 | 12 | L |
| MD72 | 19 | 13 | 20 | 30 | 16 | 11 | 26 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| MD73 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 14 | 10 | 10 | 16 | 17 | 23 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 15 | 12 | G2a |
| MD74 | 18 | 13 | 19 | 29 | 17 | 10 | 25 | 14 | 11 | 11 | 16 | 17 | 25 | 23 | 11 | 10 | 9 | 13 | 18 | 13 | 18 | 15 | 12 | L |
| YS01 | 16 | 13 | 21 | 30 | 16 | 11 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| YS02 | 18 | 13 | 20 | 30 |  | 11 | 23 | 13 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 13 | R1a |
| YS03 | 17 | 13 | 20 | 30 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS04 | 18 | 13 | 20 | 30 |  | 10 | 22 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 13 | R1a |
| YS05 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 17 | 26 | 24 | 11 | 11 | 10 | 13 | 15 | 11 | 13 | 15 | 13 | R1a |
| YS06 | 17 | 13 | 20 | 31 | 16 | 10 | 23 | 12 | 12 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 8 | 13 | 15 | 11 | 14 | 16 | 13 | R1a |
| YS07 | 18 | 14 | 19 | 31 | 17 | 10 | 26 | 11 | 11 | 10 | 15 | 15 | 22 | 21 | 12 | 11 | 11 | 13 | 15 | 12 | 21 | 16 | 12 | L |
| YS08 | 18 | 13 | 20 | 29 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 23 | 24 | 11 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS09 | 18 | 13 | 19 | 28 | 14 | 10 | 22 | 13 | 12 | 12 | 15 | 17 | 23 | 24 | 13 | 14 | 10 | 12 | 17 | 11 | 14 | 16 | 12 | R1b |

| | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| YS10 | 18 | 13 | 20 | 29 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 23 | 24 | 11 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS11 | 16 | 13 | 20 | 29 | 15 | 10 | 24 | 11 | 12 | 10 | 16 | 14 | 24 | 22 | 12 | 14 | 13 | 13 | 17 | 9 | 17 | 15 | 11 | L |
| YS12 | 19 | 14 | 19 | 30 | 14 | 11 | 26 | 13 | 11 | 11 | 14 | 18 | 24 | 23 | 11 | 10 | 10 | 14 | 19 | 13 | 19 | 14 | 12 | Q |
| YS13 | 18 | 13 | 20 | 29 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 23 | 24 | 11 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS14 | 16 | 12 | 19 | 28 | 14 | 11 | 23 | 13 | 12 | 10 | 15 | 14 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| YS15 | 20 | 14 | 20 | 31 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS16 | 18 | 13 | 20 | 31 | 17 | 11 | 23 | 13 | 12 | 11 | 14 | 20 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| YS17 | 18 | 13 | 20 | 31 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 25 | 10 | 12 | 11 | 13 | 16 | 11 | 14 | 15 | 12 | R1a |
| YS18 | 19 | 12 | 21 | 31 | 14 | 11 | 23 | 12 | 11 | 10 | 16 | 20 | 23 | 22 | 11 | 11 | 9 | 13 | 16 | 14 | 16 | 15 | 12 | G2a |
| YS19 | 19 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS20 | 17 | 14 | 20 | 32 | 15 | 10 | 23 | 12 | 13 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| YS21 | 21 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | G2a |
| YS22 | 17 | 13 | 20 | 30 | 16 | 10 | 22 | 11 | 10 | 10 | 16 | 19 | 21 | 22 | 12 | 11 | 11 | 13 | 17 | 13 | 14 | 15 | 14 | R1a |
| YS23 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 13 | 14 | 11 | 14 | 15 | 12 | R1a |
| YS24 | 19 | 13 | 20 | 30 | 16 | 11 | 24 | 13 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| YS25 | 16 | 12 | 19 | 28 | 14 | 10 | 23 | 13 | 12 | 11 | 15 | 15 | 25 | 23 | 13 | 15 | 10 | 11 | 14 | 13 | 16 | 16 | 12 | L |
| YS26 | 18 | 14 | 18 | 30 | 14 | 10 | 23 | 12 | 13 | 12 | 15 | 18 | 25 | 21 | 12 | 15 | 13 | 13 | 17 | 12 | 15 | 15 | 10 | R1b |
| YS27 | 18 | 13 | 20 | 31 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 23 | 25 | 11 | 11 | 10 | 15 | 16 | 11 | 14 | 15 | 12 | R1a |
| YS28 | 20 | 13 | 20 | 29 | 14 | 11 | 22 | 11 | 12 | 12 | 15 | 17 | 23 | 24 | 13 | 14 | 10 | 11 | 16 | 11 | 14 | 17 | 12 | R1b |
| YS29 | 17 | 13 | 0 | 30 | 14 | 10 | 24 | 12 | 11 | 10 | 14 | 17 | 18 | 23 | 12 | 11 | 11 | 14 | 17 | 14 | 15 | 15 | 12 | E1b1a |
| YS30 | 17 | 14 | 20 | 32 | 16 | 10 | 23 | 12 | 13 | 11 | 14 | 18 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 15 | 15 | 11 | R1a |
| YS31 | 18 | 13 | 19 | 29 | 15 | 10 | 27 | 14 | 11 | 9 | 15 | 19 | 21 | 24 | 11 | 11 | 11 | 13 | 16 | 14 | 16 | 14 | 12 | I2a (xI2a1) |
| YS32 | 16 | 14 | 21 | 31 | 16 | 10 | 19 | 13 | 10 | 10 | 15 | 16 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| YS33 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 17 | 26 | 24 | 11 | 11 | 10 | 13 | 15 | 11 | 13 | 15 | 13 | R1a |
| YS34 | 16 | 13 | 21 | 30 | 15 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 23 | 23 | 11 | 11 | 12 | 13 | 20 | 13 | 16 | 16 | 12 | G2a |
| YS35 | 18 | 13 | 20 | 29 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| YS36 | 19 | 14 | 21 | 32 | 15 | 10 | 24 | 12 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 17 | 11 | 15 | 15 | 12 | R1a |
| YS37 | 17 | 12 | 18 | 30 | 15 | 10 | 26 | 13 | 12 | 11 | 14 | 18 | 23 | 24 | 11 | 11 | 11 | 14 | 16 | 13 | 16 | 15 | 12 | Q |
| YS38 | 16 | 14 | 19 | 30 | 13 | 10 | 23 | 13 | 10 | 9 | 14 | 14 | 20 | 23 | 11 | 13 | 11 | 14 | 15 | 13 | 16 | 14 | 12 | L |
| YS39 | 18 | 12 | 19 | 28 | 14 | 10 | 23 | 12 | 12 | 10 | 15 | 15 | 25 | 22 | 12 | 14 | 10 | 11 | 15 | 13 | 17 | 15 | 12 | L |
| YS40 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 17 | 26 | 24 | 11 | 11 | 10 | 13 | 15 | 11 | 13 | 15 | 13 | R1a |
| YS41 | 17 | 12 | 19 | 28 | 16 | 10 | 23 | 13 | 11 | 11 | 16 | 18 | 24 | 23 | 10 | 10 | 10 | 14 | 19 | 13 | 17 | 15 | 12 | L |
| YS42 | 18 | 13 | 20 | 30 | 16 | 10 | 22 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 16 | 13 | R1a |
| YS43 | 17 | 13 | 21 | 30 | 14 | 10 | 24 | 12 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 10 | 13 | 15 | 15 | 20 | 15 | 13 | E1b1a |
| YS44 | 19 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| YS45 | 18 | 13 | 20 | 30 | | 10 | 23 | 12 | 12 | 11 | 14 | 18 | 24 | 25 | 10 | 11 | 8 | 13 | 16 | 11 | 14 | 16 | 13 | R1a |
| YS46 | 20 | 13 | 20 | 30 | | 11 | 25 | 12 | 12 | 11 | 0 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS47 | 19 | 14 | 21 | 32 | 15 | 10 | 24 | 12 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 11 | 13 | 17 | 11 | 14 | 15 | 12 | R1a |
| YS48 | 17 | 13 | 19 | 32 | 15 | 10 | 21 | 11 | 12 | 10 | 16 | 16 | 21 | 23 | 13 | 11 | 15 | 14 | 16 | 14 | 15 | 17 | 12 | G2A |
| YS49 | 18 | 13 | 20 | 30 | 16 | 10 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 17 | 11 | 14 | 15 | 11 | R1a |
| YS50 | 18 | 13 | 20 | 30 | | 10 | 23 | 13 | 12 | 11 | 14 | 18 | 24 | 25 | 10 | 11 | 8 | 13 | 16 | 11 | 14 | 16 | 13 | R1a |
| YS51 | 15 | 13 | 20 | 29 | 15 | 10 | 24 | 12 | 11 | 10 | 16 | 15 | 22 | 20 | 12 | 14 | 11 | 12 | 18 | 9 | 16 | 15 | 12 | L |
| YS52 | 16 | 13 | 21 | 30 | 15 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 23 | 23 | 11 | 11 | 12 | 13 | 20 | 13 | 16 | 16 | 12 | G2a |
| YS53 | 17 | 12 | 18 | 28 | 15 | 10 | 26 | 13 | 12 | 11 | 14 | 18 | 23 | 24 | 11 | 14 | 11 | 14 | 16 | 13 | 16 | 15 | 11 | Q |
| YS54 | 17 | 14 | 19 | 30 | 15 | 10 | 26 | 12 | 11 | 11 | 16 | 17 | 25 | 23 | 11 | 10 | 10 | 14 | 19 | 14 | 18 | 16 | 13 | Q |
| YS55 | 18 | 13 | 20 | 29 | 17 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS56 | 20 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS57 | 20 | 13 | 20 | 30 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 14 | 11 | 14 | 15 | 12 | R1a |
| YS58 | 18 | 14 | 20 | 30 | 15 | 10 | 24 | 12 | 12 | 11 | 14 | 16 | 23 | 25 | 10 | 11 | 10 | 13 | 17 | 11 | 16 | 15 | | R1a |
| YS59 | 19 | 13 | 21 | 32 | 15 | 10 | 24 | 12 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 17 | 11 | 14 | 15 | 12 | R1a |
| YS60 | 16 | 12 | 19 | 28 | 14 | 11 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 12 | 18 | 16 | 12 | L |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| YS61 | 19 | 14 | 20 | 30 | 14 | 10 | 22 | 11 | 12 | 10 | 14 | 18 | 23 | 23 | 12 | 13 | 11 | 12 | 17 | 15 | 17 | 14 | 12 | L |
| YS62 | 17 | 12 | 19 | 28 | 15 | 10 | 25 | 14 | 11 | 11 | 16 | 19 | 24 | 23 | 10 | 10 | 10 | 14 | 19 | 13 | 16 | 15 | 12 | L |
| YS63 | 16 | 13 | 20 | 29 | 15 | 10 | 24 | 11 | 12 | 10 | 16 | 14 | 24 | 22 | 12 | 14 | 13 | 13 | 17 | 9 | 17 | 15 | 11 | L |
| YS64 | 18 | 14 | 20 | 31 | 14 | 10 | 25 | 13 | 11 | 10 | 14 | 16 | 21 | 24 | 12 | 11 | 9 | 12 | 18 | 11 | 16 | 15 | 11 | J1 |
| YS65 | 17 | 13 | ? | 29 | 14 | 10 | 24 | 12 | 11 | 10 | 14 | 17 | 18 | 23 | 12 | 11 | 11 | 14 | 17 | 14 | 15 | 15 | 12 | E1b1a |
| YS66 | 18 | 13 | 20 | 29 | ? | 11 | 23 | 12 | 12 | 11 | 14 | 19 | ? | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| YS67 | 17 | 14 | 19 | 30 | 15 | 10 | 24 | 15 | 11 | 11 | 15 | 15 | 25 | 23 | 11 | 10 | 10 | 14 | 18 | 13 | 17 | 15 | 12 | L |
| YS68 | 19 | 14 | 21 | 32 | 15 | 10 | 24 | 12 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 17 | 11 | 14 | 15 | 13 | R1a |
| YS69 | 19 | 13 | 20 | 32 | 15 | 10 | 23 | 13 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| YS70 | 18 | 14 | 20 | 33 | 15 | 10 | 23 | 12 | 12 | 10 | 14 | 20 | 23 | 25 | 10 | 12 | 10 | 13 | 14 | 11 | 14 | 16 | 12 | R1a |
| YS71 | 16 | 13 | 21 | 29 | 17 | 11 | 19 | 13 | 10 | 10 | 16 | 18 | 25 | 23 | 11 | 11 | 12 | 13 | 17 | 13 | 16 | 15 | 11 | G2a |
| YS72 | 19 | 13 | 20 | 30 | 17 | 11 | 23 | 12 | 11 | 11 | 14 | 19 | 24 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| YS73 | 17 | 13 | 22 | 30 | 14 | 10 | 24 | 12 | 11 | 10 | 14 | 17 | 18 | 23 | 12 | 11 | 11 | 14 | 17 | 14 | 15 | 15 | 12 | E1b1a |
| YS74 | 18 | 13 | 20 | 31 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 25 | 10 | 12 | 11 | 13 | 16 | 11 | 14 | 15 | 13 | R1a |
| YS75 | 10 | 14 | 20 | 31 | 15 | 11 | 25 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM01 | 18 | 13 | 20 | 30 | 16 | 10 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM02 | 16 | 12 | 19 | 29 | 15 | 11 | 23 | 12 | 9 | 9 | 15 | 15 | 21 | 24 | 11 | 11 | 9 | 13 | 16 | 13 | 18 | 13 | 11 | J2b |
| KM03 | 17 | 13 | 19 | 20 | 15 | 9 | 23 | 12 | 12 | 9 | 15 | 15 | 21 | 23 | 11 | 11 | 11 | 13 | 15 | 14 | 19 | 15 | 12 | L |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KM04 | 18 | 13 | 21 | 29 | 15 | 10 | 23 | 11 | 12 | 9 | 15 | 16 | 21 | 23 | 12 | 11 | 10 | 12 | 15 | 15 | 18 | 15 | 12 | L |
| KM05 | 18 | 14 | 19 | 31 | 15 | 10 | 24 | 13 | 11 | 11 | 15 | 19 | 26 | 23 | 10 | 10 | 10 | 14 | 18 | 13 | 19 | 15 | 11 | Q |
| KM06 | 16 | 13 | 19 | 29 | 15 | 10 | 23 | 13 | 12 | 10 | 14 | 19 | 20 | 22 | 11 | 11 | 9 | 12 | 17 | 14 | 18 | 15 | 11 | H |
| KM08 | 16 | 13 | 19 | 29 | 15 | 10 | 24 | 11 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM07 | 17 | 14 | 20 | 30 | 16 | 10 | 24 | 11 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM09 | 18 | 13 | 20 | 30 | 15 | 12 | 24 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM10 | 18 | 13 | 20 | 31 | 16 | 11 | 23 | 11 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 16 | 10 | 14 | 17 | 12 | R1a |
| KM11 | 18 | 13 | 20 | 30 | 16 | 11 | 24 | 12 | 12 | 11 | 14 | 21 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KM12 | 20 | 14 | 20 | 31 | 16 | 10 | 24 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 17 | 12 | 14 | 15 | 11 | R1a |
| KM13 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 25 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| KM14 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM15 | 18 | 13 | 20 | 29 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 21 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM16 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 24 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| KM17 | 17 | 13 | 19 | 29 | 16 | 10 | 25 | 11 | 12 | 10 | 16 | 16 | 21 | 22 | 13 | 16 | 11 | 12 | 17 | 9 | 16 | 14 | 12 | L |
| KM18 | 18 | 13 | 20 | 30 | 16 | 10 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM19 | 17 | 13 | 19 | 29 | 16 | 10 | 25 | 11 | 12 | 10 | 16 | 16 | 21 | 22 | 13 | 16 | 11 | 12 | 17 | 9 | 16 | 14 | 12 | L |
| KM20 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM21 | 18 | 13 | 20 | 30 | 16 | 11 | 24 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM22 | 18 | 13 | 20 | 30 | 16 | 11 | 24 | 12 | 12 | 11 | 14 | 21 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM23 | 18 | 13 | 20 | 30 | 16 | 10 | 24 | 13 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 9 | 13 | 16 | 14 | 20 | 16 | 11 | H |
| KM24 | 18 | 13 | 20 | 30 | 16 | 11 | 24 | 12 | 12 | 11 | 14 | 21 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM25 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM26 | 18 | 13 | 20 | 30 | 16 | 10 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM27 | 16 | 13 | 19 | 29 | 16 | 9 | 24 | 12 | 11 | 10 | 14 | 16 | 18 | 22 | 10 | 11 | 12 | 15 | 18 | 13 | 20 | 15 | 10 | E1b1a |
| KM28 | 14 | 12 | 18 | 28 | 15 | 10 | 23 | 12 | 11 | 9 | 15 |  | 21 | 24 | 12 | 11 | 9 | 12 | 15 | 14 | 17 | 13 | 11 | J2b |
| KM29 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM30 | 18 | 14 | 20 | 31 | 14 | 10 | 23 | 13 | 12 | 11 | 14 | 19 | 24 | 25 | 10 | 11 | 10 | 13 | 16 | 11 | 14 | 15 | 11 | R1a |
| KM31 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 14 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM32 | 18 | 13 | 20 | 30 | 16 | 11 | 24 | 12 | 12 | 11 | 14 | 20 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM33 | 18 | 13 | 20 | 30 | 16 | 10 | 24 | 13 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 9 | 13 | 16 | 13 | 19 | 16 | 11 | J1 |
| KM34 | 18 | 13 | 20 | 30 | 0 | 10 | 24 | 13 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 9 | 13 | 16 | 14 | 20 | 16 | 11 | H |
| KM35 | 18 | 13 | 20 | 30 | 0 | 10 | 24 | 13 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 9 | 13 | 16 | 14 | 20 | 16 | 11 | H |
| KM36 | 19 | 13 | 20 | 30 | 0 | 11 | 23 | 13 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| KM37 | 16 | 12 | 19 | 28 | 14 | 10 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KM38 | 17 | 12 | 18 | 28 | 0 | 9 | 24 | 13 | 12 | 11 | 16 | 17 | 24 | 24 | 12 | 10 | 11 | 14 | 17 | 14 | 20 | 15 | 12 | Q |
| KM39 | 18 | 13 | 20 | 30 | 14 | 10 | 24 | 13 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 9 | 13 | 16 | 14 | 20 | 16 | 11 | H |
| KM40 | 15 | 13 | 20 | 29 | 14 | 10 | 24 | 12 | 11 | 9 | 16 | 17 | 21 | 23 | 12 | 11 | 10 | 12 | 17 | 13 | 17 | 15 | 11 | J1 |
| KM41 | 18 | 13 | 20 | 30 | 15 | 10 | 24 | 13 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 9 | 13 | 16 | 14 | 20 | 16 | 11 | H |
| KM42 | 18 | 13 | 20 | 30 | 16 | 10 | 24 | 13 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 9 | 13 | 16 | 14 | 20 | 16 | 11 | H |
| KM43 | 18 | 13 | 20 | 30 | 14 | 10 | 24 | 13 | 11 | 9 | 14 | 15 | 21 | 23 | 11 | 11 | 9 | 13 | 16 | 14 | 20 | 16 | 11 | H |
| KM44 | 16 | 12 | 19 | 24 | 14 | 10 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM45 | 16 | 13 | 21 | 30 | 16 | 10 | 19 | 13 | 10 | 10 | 16 | 18 | 23 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| KM46 | 18 | 13 | 21 | 30 | 14 | 10 | 23 | 13 | 11 | 9 | 15 | 16 | 22 | 23 | 13 | 11 | 10 | 12 | 15 | 13 | 15 | 15 | 12 | J1 |
| KM47 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| KM48 | 17 | 12 | 18 | 28 | 14 | 9 | 26 | 13 | 12 | 11 | 16 | 17 | 24 | 24 | 12 | 10 | 11 | 15 | 17 | 14 | 20 | 15 | 12 | Q |
| KM49 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | R1a |
| KM50 | 17 | 12 | 18 | 28 | 0 | 9 | 26 | 13 | 12 | 11 | 16 | 17 | 24 | 24 | 12 | 10 | 11 | 15 | 17 | 14 | 20 | 15 | 12 | Q |
| KM51 | 16 | 12 | 19 | 28 | 14 | 10 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 12 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM52 | 18 | 13 | 20 | 30 | 15 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 12 | Q |
| KM53 | 16 | 12 | 19 | 26 | 14 | 10 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM54 | 17 | 12 | 18 | 28 | 14 | 9 | 26 | 13 | 12 | 11 | 16 | 17 | 24 | 24 | 12 | 10 | 11 | 15 | 17 | 14 | 20 | 15 | 12 | Q |
| KM55 | 19 | 13 | 20 | 30 | 16 | 11 | 23 | 13 | 12 | 11 | 14 | 18 | 24 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 13 | R1a |
| KM56 | 17 | 12 | 18 | 28 | 14 | 9 | 26 | 13 | 12 | 11 | 16 | 17 | 24 | 24 | 12 | 10 | 11 | 15 | 17 | 14 | 20 | 15 | 12 | Q |
| KM57 | 17 | 12 | 18 | 28 | 14 | 9 | 26 | 13 | 12 | 11 | 16 | 17 | 24 | 24 | 12 | 10 | 11 | 14 | 17 | 14 | 20 | 15 | 12 | Q |
| KM58 | 16 | 12 | 19 | 28 | 14 | 11 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM59 | 16 | 12 | 19 | 28 | 14 | 10 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM60 | 18 | 13 | 19 | 28 | 14 | 10 | 22 | 13 | 12 | 12 | 15 | 17 | 23 | 24 | 13 | 14 | 10 | 12 | 17 | 11 | 14 | 16 | 12 | R1b |
| KM61 | 16 | 13 | 21 | 30 | 17 | 10 | 19 | 13 | 10 | 10 | 16 | 17 | 23 | 23 | 11 | 11 | 12 | 13 | 19 | 13 | 16 | 16 | 12 | G2a |
| KM62 | 17 | 12 | 18 | 28 | 14 | 9 | 26 | 13 | 12 | 11 | 16 | 17 | 24 | 24 | 12 | 10 | 11 | 14 | 17 | 14 | 19 | 15 | 12 | Q |
| KM63 | 16 | 12 | 19 | 28 | 14 | 11 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM64 | 18 | 13 | 21 | 30 | 14 | 10 | 23 | 13 | 11 | 9 | 15 | 16 | 22 | 23 | 13 | 11 | 10 | 12 | 15 | 13 | 15 | 15 | 12 | J1 |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KM65 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 19 | 23 | 25 | 10 | 11 | 10 | 13 | 14 | 11 | 14 | 15 | 12 | R1a |
| KM66 | 18 | 13 | 21 | 30 | 14 | 10 | 23 | 13 | 11 | 9 | 15 | 16 | 22 | 23 | 13 | 11 | 10 | 12 | 15 | 13 | 15 | 15 | 12 | J1 |
| KM67 | 17 | 12 | 18 | 28 | 14 | 9 | 26 | 13 | 12 | 11 | 16 | 17 | 24 | 24 | 12 | 10 | 11 | 15 | 17 | 14 | 20 | 15 | 12 | Q |
| KM68 | 18 | 13 | 20 | 30 | 16 | 11 | 23 | 12 | 12 | 11 | 14 | 18 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 14 | 15 | 11 | R1a |
| KM69 | 19 | 14 | 20 | 31 | 16 | 10 | 24 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 17 | 12 | 14 | 15 | 11 | R1a |
| KM70 | 19 | 13 | 20 | 30 | 15 | 10 | 26 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 14 | 11 | 14 | 15 | 12 | R1a |

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KM71 | 19 | 14 | 20 | 31 | 16 | 10 | 24 | 12 | 12 | 11 | 14 | 19 | 24 | 24 | 10 | 11 | 10 | 13 | 17 | 12 | 14 | 15 | 11 | R1a |
| KM72 | 16 | 12 | 19 | 28 | 14 | 11 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM73 | 16 | 12 | 19 | 28 | 14 | 11 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM74 | 16 | 12 | 19 | 28 | 14 | 11 | 23 | 13 | 12 | 10 | 15 | 15 | 23 | 22 | 13 | 14 | 10 | 11 | 15 | 13 | 19 | 16 | 12 | L |
| KM75 | 20 | 13 | 20 | 30 | 15 | 11 | 26 | 12 | 12 | 11 | 14 | 19 | 23 | 24 | 10 | 11 | 10 | 13 | 15 | 11 | 13 | 15 | 12 | R1a |